7/25/68

# THE DEVELOPMENT OF A SELF-ADAPTIVE PREDICTION AND CONTROL SYSTEM

A THESIS

Presented to

The Faculty of the Division of Graduate

Studies and Research

by

David Edward Ferguson

In Partial Fulfillment

of the Requirements for the Degree

Master of Science

in the School of Industrial and Systems Engineering

Georgia Institute of Technology

May, 1971

THE DEVELOPMENT OF A SELF-ADAPTIVE

PREDICTION AND CONTROL SYSTEM

Approved:

Chairman ______________________

______________________

______________________

Date approved by Chairman: May 28, 1971

# ACKNOWLEDGMENTS

The author was encouraged to pursue the development of this self-adaptive prediction and control system by Dr. Douglas C. Montgomery, who also served as the thesis advisor. My special thanks go to him for his suggestions, patience and support. Credit is also given to Dr. Jack R. Walker and to Dr. Lynwood A. Johnson, both of whom contributed time and constructive comments.

My warmest personal appreciation goes to my wife, Lucy, whose patience and understanding made this moment possible for me.

TABLE OF CONTENTS

# LIST OF TABLES

## SUMMARY

This thesis is concerned with the development of a self-adaptive prediction and control model for use whenever it is necessary to predict or control the value of some time series of interest. By being self-adaptive, it is meant that the model automatically adjusts itself so as to compensate rapidly for changes in the basic nature of the time series being predicted. The well-known Box-Jenkins prediction and control model was selected as the basis for the research. This model was then made self-adaptive by the application of evolutionary operation (EVOP) techniques. Two EVOP techniques, the standard factorial design EVOP procedure and the sequential application of the simplex design, were used to develop two alternative self-adaptive versions of the basic Box-Jenkins model. These two self-adaptive versions were compared both with each other and with the basic model, using prediction accuracy measures, to determine if an improved prediction and control model resulted.

Seven time series possessing certain patterns and random components were generated. The three forms of the model, the basic Box-Jenkins model and its two EVOP modifications, were used to predict these seven series. Their various performances were compared using the relative accuracies of the predictions as the measure of effectiveness of the models.

The conclusions reached concerning the application of EVOP procedures to the basic Box-Jenkins model are as follows:

1. The application of the sequential simplex EVOP procedures yields a predictive performance record for the seven time series tested that is no better than that obtained from the application of the basic Box-Jenkins model.

2. The application of the factorial design EVOP procedures yields a predictive performance record that is significantly better than that obtained from the application of the basic Box-Jenkins model to the seven time series.

# CHAPTER I

## INTRODUCTION

*If we could first know where we are and whither we are tending, we could better judge what to do and how to do it.*†

This quotation summarizes the problems of prediction and control to be investigated in this study in that it is our objective to determine not only "where we are and whither we are tending" but also determine "what to do and how to do it." These two problems, those of prediction and control, are of special interest to managers, salesmen, engineers, and others concerned with the quality or quantity of industrial products. A brief description of these problems will reveal their importance to the successful operation of a manufacturing enterprise and indicate the characteristics that solutions to these problems must possess.

The problem of control suggests that it is desired to operate some process--chemical, mechanical, electrical, etc.--in such a way as to attain the best value of some objective function such as yield, cost or performance. This objective function is determined by the values of certain variables, some of which are controllable and some of which are not. If values of the controllable variables can be determined which yield the optimum value of the objective function, then the process

---

†Abraham Lincoln as quoted by Spencer (14).

is said to be under control. However, the presence of uncontrollable variables implies that it may be difficult or impossible to determine the best values of the controllable variables. This may happen because these best values may vary from one time period to another under the influence of the uncontrollable variables. In other words both the process and the controllable variables become stochastic in nature. However, it is still feasible to consider controlling a stochastic process. Stochastic control consists of estimating the values of the controllable variables so that the resulting value of the objective function is as close as is possible to the true optimum value. If the deviation between the true optimum value of the objective function and the actual obtained value can be minimized by properly choosing the values of the controllable variables, then the stochastic process can be controlled.

The problem of prediction suggests that it is desired to know in advance the future value of some quantity, such as a process characteristic or the demand for some inventory item. In problems of prediction there may or may not be variables that can be adjusted to influence the future value of the quantity of interest. In many instances, however, it is assumed that in problems of prediction the underlying process generating the individual quantities will continue to operate regardless of any possible adjustments that may be made to relevant variables. It is therefore not required to be able to control the future values of the quantity of interest. It is only important to determine or predict as accurately as is possible what the future values of the quantity of interest will be. For example, if a manager knows in advance the quantity of a given item that will be demanded in a coming time period, he can better adjust his inventories. Intuitively it seems that solutions to these two problems of prediction and control are related. Indeed this is true and in a succeeding chapter of this thesis a mathematical relationship between the two problems will be shown.

## Scope of the Thesis

This thesis is concerned with the development of a solution procedure for the control and prediction problems. This solution procedure will build upon a prediction and control model developed by Box and Jenkins (4). Details of this model will be given in a succeeding chapter. The proposed solution procedure, like the Box-Jenkins model, operates using historical values of a time series as input. This time series consists either of the historical values of a controllable variable or of historical values of some quantity to be predicted. Models using such a simple input are in many cases more advantageous than those requiring complex input data and functional relationships. Motivations for the use of such a model include:

(1) a lack of understanding of the basic theory causing the time series;

(2) the inability to justify a more sophisticated approach economically; or

(3) the insufficiency of time to pursue other approaches.

## Objective of the Thesis

The objective of the thesis is to develop a model for the solution of the prediction and control problems which will be an improvement to the existing Box-Jenkins procedure. The new model will use the existing Box-Jenkins model as a basis; however, it will be self-adaptive in the sense that the parameters in the model are automatically adjusted to compensate for changes in the basic nature of the time series. Two evolutionary operation techniques will be used to make the basic Box-Jenkins model self-adaptive. Results obtained by operating both the basic Box-Jenkins model and the two revised models will be compared to determine:

(1) if applying evolutionary operation techniques to the Box-Jenkins model results in improved prediction and control; and

(2) which evolutionary operation technique is the better to use in the event that an improved solution does result.

## Other Adaptive Prediction Systems

Several examples of adaptive prediction or forecasting systems are found in the literature. These systems are predominantly used in a sales or product demand forecasting context and involve the use of models other than the Box-Jenkins model. However, there is similarity in that the adaptive nature of these systems is based on the alteration or variation of parameters in the system in response to changes in the nature of the time series of interest. These parameters are usually exponential smoothing parameters as opposed to the control parameters in the Box-Jenkins model. However, even though the basic parameters are somewhat different, it is of value to review briefly some of these adaptive systems as a prelude to the development of the system in this thesis.

Chow (9) has developed an adaptive forecasting system based on a simple exponential smoothing model involving a linear trend correction. However, instead of the customary use of a single, constant exponential smoothing parameter, he proposes that three different smoothing parameters be used to forecast three different values of the time series for each period. The three constants are set at high, normal and low levels. The forecast that is used initially is the one resulting from the normal value of the smoothing constant. However, when one of the outer forecasts yields a smaller forecast error than the normal forecast, then the outer smoothing constant value is established as the normal value for the coming time period. High and low constant values are re-established about the new normal value and the process is repeated.

Brown (5) proposes the use of tracking signals to monitor the performance of the forecasting system being used whether it is a simple exponential smoothing model or one involving trend and seasonal corrections. If the tracking signal indicates that the forecasting system is out of control, Brown recommends taking such corrective action as increasing or decreasing the smoothing parameters as necessary. Such procedures, however, need administrative attention and can be quite expensive and time consuming.

Burgess (7) presents a tracking signal monitoring system similar to Brown's to signal out of control conditions. However, he presents a procedure for automatically adjusting the smoothing parameter whenever an out of control condition is indicated. The smoothing parameter, $\alpha$, is defined as $\frac{1}{1+M}$ where M is the number of time periods to the mid-point of an exponentially smoothed moving average. For each period in which the tracking signal indicates that the forecasting system is in control, M is incremented by a value of one until the maximum M value of 20 is reached. This maximum M value corresponds to a minimum alpha value of 0.05. Small values of $\alpha$ should be used with data if we have confidence that the future time series behavior will be similar to past behavior. However, if the time series begins to change significantly, and the tracking signal indicates an out of control condition, the following action is automatically taken. A constant value MS is subtracted from the current value of M for each period in which the out of control condition is indicated. This continues until a minimum M value is reached corresponding to a maximum alpha value. These larger alpha values seem intuitively justifiable since in a changing time series more recent data should be used on which to base a prediction. High exponential smoothing constants weight recent data more heavily than older data.

Roberts (13) has developed an adaptive forecasting system based on the forecasting models of Winters (17). He used an evolutionary operation procedure involving a factorial experimental design to test the effects on forecast accuracy of varying the exponential smoothing parameters in the forecasting model. If an effect of varying one or more of the smoothing parameters is statistically significant, then the center point of the experimental design is shifted or altered accordingly. Successive forecasts are made for each period using the parameter combination defining the center point of the experimental design. Whenever the center point is moved, the forecast is adapted in response. Roberts has shown his system to be superior to those of Brown, Winters and Chow with respect to certain accuracy and response criteria.

Montgomery (11) has also used an evolutionary operation scheme to make a forecasting system adaptive in nature. The evolutionary operation procedure involves the use of a different experimental design from those used by Roberts and Chow. Montgomery recommends the use of the simplex, which like the factorial, is an orthogonal, first order experimental design. This procedure involves changing the exponential smoothing parameters each period by the sequential application of the simplex design. A new simplex is formed each period by deleting only one point from the previous simplex and adding one new point as defined by fixed relationships. The point that is deleted each period is the parameter combination which yields the forecast resulting in the largest forecast error. Thus the design theoretically insures that the forecasting system will traverse the parameter space from points of high forecast error to points of lower forecast error. As yet no comparisons have been made directly between the work of Roberts and that of Montgomery, although Montgomery has shown his procedure to be superior to that of Chow.

# CHAPTER II

## THE BOX-JENKINS PREDICTION AND CONTROL MODEL

In this chapter the Box-Jenkins prediction and control model will be presented and discussed. This model was first presented in a paper read at a Research Methods Meeting of the Royal Statistical Society on April 4, 1962 (4). The model uses empirical feedback as opposed to technical feedback in its operation. Empirical feedback occurs when it is possible to state a simple rule which describes what action should be taken and what new experiments should be done in every conceivable situation. On the other hand, technical feedback occurs when the information coming from the experiment interacts with technical knowledge contained in the experimenter's mind to lead to some form of action. Box and Jenkins explain the concept of adaptive optimization in terms of its being related to empirical feedback; they then develop a discrete-time adaptive optimization model. Drawing upon this discrete-time adaptive optimization model, Box and Jenkins define the general control and prediction problems and develop a formal prediction and control model. However, this model is quite complicated and computationally infeasible in actual practice; therefore, Box and Jenkins develop a more practical prediction and control model, based on the formal model, but incorporating generalizations and simplifications that make it practical.

This research is concerned with improving the practical prediction model by making the control parameters in the model self-adaptive to changes in the basic nature of the time series of interest. Therefore, this chapter is not concerned with the formal model; rather, the more practical model will be discussed with reference to the formal model and the concepts of adaptive optimization as necessary. It should be noted that Box and Jenkins give conditions under which, and define the stochastic processes for which, both their formal and practical models yield optimal solutions. However, in the time series to be used in this thesis, these conditions are not attained or adhered to. Therefore, no claims of optimality can be made for the solutions obtained from the Box-Jenkins model in this thesis. This is true of both the basic Box-Jenkins model and of the model after the parameters in it have been made self-adaptive.

## Discrete Adaptive Optimization

To discuss the Box-Jenkins model it is first necessary to consider some of the concepts and definitions of discrete adaptive optimization. In a process such as a chemical or manufacturing process, there are usually certain variables of interest that can be effectively controlled. These variables can be controlled either directly, such as pressure which can be controlled by opening or closing a valve, or indirectly such as process yield which may be controlled by adjusting a variable several stages upstream in the production process. For the sake of simplicity in this presentation, it will be assumed that there is a single controllable variable of interest, X, which can be controlled directly. Box and Jenkins (4) present the analogous development when the variable X can be controlled only indirectly by adjusting some manipulated variable $X^*$. It is also assumed that data concerning the controllable variable are available only at discrete and equal intervals of time, each of which is called a phase. The physical situation usually remains constant during a given phase but may change from phase to phase.

Inherent in the process under consideration are uncontrollable and unmeasurable variables having levels $\varepsilon_p$ during the pth phase and changing from phase to phase. If $\eta(X)$ is the response of the process in relation to the controllable variables, then $\eta(X|\varepsilon_p)$ is the conditional response in relation to the controllable variables given $\varepsilon$, the uncontrollable variables. Suppose that in the pth phase this response function may be approximated by the quadratic equation

$$\eta_p = \eta(X|\varepsilon_p) = \eta(\theta_p) - \frac{1}{2}\beta_{11}(X-\theta_p)^2$$

where $\theta_p = (X_{max}|\varepsilon_p)$ is the conditional optimum setting of the controllable variable during the pth phase; and $\beta_{11}$ is a constant known from prior calibration which does not change appreciably with $\varepsilon_p$. Because of the presence of the uncontrollable variable $\varepsilon_p$, $\theta_p$ can be assumed to follow some non-stationary stochastic process. If $X_p$ is the set-point at which the controllable variable is set in the pth phase, then the standardized slope of the response function evaluated at $X = X_p$ is

$$\frac{1}{B_{11}}\left(\frac{d\eta_p}{dX}\right)_{X=X_p} = \theta_p - X_p = \epsilon_p.$$

It can be easily seen that $\epsilon_p$ measures the extent to which $X_p$ deviates from the optimal value $\theta_p$. If further experiments are performed about $X_p$, say at $X_p + \delta$ and $X_p - \delta$, then the average response observed at these levels is $\bar{Y}(X_p+\delta)$ and $\bar{Y}(X_p-\delta)$. Thus an estimate $e_p$ of $\epsilon_p = \theta_p - X_p$ is given by

$$e_p = \frac{\bar{Y}(X_p+\delta) - \bar{Y}(X_p-\delta)}{2(\delta\beta_{11})} = \epsilon_p + u_p$$

where $u_p$ is a measurement error. It is easy to write $e_p = \epsilon_p + u_p = \theta_p - X_p + u_p = Z_p - X_p$ where $Z_p = \theta_p + u_p$ is an estimate of the position of the optimal setting $\theta_p$ during the pth phase.

If over a period of time a series of adjustments to the process have been made on some basis, an empirical record of both the set-points $X_p, X_{p-1}, X_{p-2}, \ldots$ and the deviations $e_p, e_{p-1}, e_{p-2}, \ldots$ should be available. From these, the series of the estimated positions of the optimal values can be calculated to give the series $Z_p, Z_{p-1}, Z_{p-2}, \ldots$ using the relationship $Z_p = X_p + e_p$. From these series it is desired to make an adjustment $x_{p+1}$ to the set-point $X_p$ so that the set-point $X_{p+1} = X_p + x_{p+1}$ will in some sense be "best" in relation to the coming unknown value of $\theta_{p+1}$.

Suppose that the objective function $\eta$ is such that the loss sustained by operating the process at the set-point $X_{p+1}$ instead of at the optimum $\theta_{p+1}$ is measured by

$$\eta(\theta_{p+1}) - \eta(X_{p+1}|\varepsilon_{p+1}) = \frac{1}{2}\beta_{11}(X_{p+1}-\theta_{p+1})^2.$$

The adjustment $x_{p+1}$ will then minimize the expected loss if it is chosen so that $E(\theta_{p+1}-X_{p+1})^2$ is minimized. Box and Jenkins show that this objective is attained if $X_{p+1}$ is set equal to $\hat{\theta}_{p+1} = f(Z_p, Z_{p-1}, \ldots)$ where $f(Z_p, Z_{p-1}, \ldots)$ is the minimum mean square predictor of $\theta_{p+1}$ based on the observations of $Z_p, Z_{p-1}, Z_{p-2}, \ldots$ . If $\hat{\theta}_{p+1}$ is assumed to be a linear function of the Z's, then this predictor $\hat{\theta}_{p+1}$ can be written as $\hat{\theta}_{p+1} = \sum_{j=0}^{\infty} a_j Z_{p-j}$ where the $a_j$'s are called the predictor weights. Therefore, at the beginning of the $(p+1)^{th}$ phase, an adjustment $x_{p+1} = \hat{\theta}_{p+1} - \hat{\theta}_p$ would be applied to the previous set-point $X_p$. However, since the e's are observed directly, the above adjustment might be written as $x_{p+1} = X_{p+1} - X_p = \hat{\theta}_{p+1} - \hat{\theta}_p = \sum_{j=0}^{\infty} b_j e_{p-j}$ where the $b_j$'s are called the controller weights. Box and Jenkins show that a definite relationship exists between the predictor weights and the controller weights. The optimal adjustment is thus obtained by choosing the a's or equivalently the b's to minimize $E(\epsilon_{p+1}^2) = E(\theta_{p+1}-\hat{\theta}_{p+1})^2$. However, $\hat{\theta}_{p+1} = f(Z_p, Z_{p-1}, Z_{p-2}, \ldots)$ and $Z_{p+1} = \theta_{p+1} + u_{p+1}$ so that if the measurement error $u_{p+1}$ is distributed about zero with variance $\sigma_u^2$ independently of $u_p, u_{p-1}, \ldots$ and of $\theta_p, \theta_{p-1}, \ldots$ then the expected error can be written as

$$E(e_{p+1}^2) = E(Z_{p+1}-\hat{\theta}_{p+1})^2 = E(\epsilon_{p+1}^2) + \sigma_u^2.$$

The loss is then minimized when $E(Z_{p+1}-\hat{\theta}_{p+1})^2 = E(e_{p+1})^2$ is minimized and $Z_{p+1}$, the best predictor of $Z_{p+1}$, gives the best predictor $\hat{\theta}_{p+1}$ of $\theta_{p+1}$ by the equation $\hat{\theta}_{p+1} = \sum_{j=0}^{\infty} a_j Z_{p-j}$.

In general $\sigma_u^2$ will depend on the number of experiments performed, n, and on $\delta$, the magnitude of the perturbations. However, for fixed values of n and $\delta$, Box and Jenkins have shown that the best way of tracking $\theta_{p+1}$ is always to make an adjustment $x_{p+1} = \hat{\theta}_{p+1} - \hat{\theta}_p = \sum_{j=0}^{\infty} b_j e_{p-j}$. If the measurement errors, u, are independent of the $\theta$'s and of each other, then $\hat{\theta}_{p+1} = \hat{Z}_{p+1}$.

## Discrete Adaptive Control

Box and Jenkins use specific concepts of adaptive optimization to develop an adaptive control model. The symbols used in this section are the same as those used previously; however, the context in which they are used is different, due to the nature of the control problem.

Assume that some quality characteristic is of interest and that if no steps were taken to control this characteristic, it would have an observed value at the pth phase of $Z_p = \theta_p + u_p$ where $\theta_p$ is the conditional optimal setting of the characteristic in the pth phase and $u_p$ is again the measurement error. It is assumed as before that because of the presence in the process of uncontrollable and unmeasurable variables that $\theta_p$ follows some non-stationary stochastic process. The objective in this control problem is to hold the value of $\theta_p$ as close as is possible to some target value. To achieve this objective, suppose that it is possible to adjust the mean value of the stochastic variable Z, the observed value of the quality characteristic, up or down. Let $-X_p$ be the total correction that has been applied at the pth phase. Thus the actually observed quantity is the apparent deviation from target $Z_p - X_p = \theta_p - X_p + u_p = \epsilon_p + u_p = e_p$ where $\epsilon_p$ and $e_p$ are the same quantities defined in the adaptive optimization problem.

It is, therefore, desired to calculate some further adjustment, $x_{p+1}$, to be made to the mean in the pth phase so that the total correction in the $(p+1)^{th}$ phase will be $-X_{p+1} = -(X_p + x_{p+1})$. It is also desired that, when this correction is applied, the actual deviation from target $\epsilon_{p+1} = \theta_{p+1} - X_{p+1}$ will be small. Assume that the loss involved by $\theta$'s being off target by an amount $\epsilon_p$ is proportional to $\epsilon_p^2$. It is thus required that $x_{p+1}$ be chosen so that $E(\epsilon_{p+1})^2 = E(\theta_{p+1} - X_{p+1})^2$ is minimized. As was shown in the section on adaptive optimization, this requires that $x_{p+1} = \hat{\theta}_{p+1} - \hat{\theta}_p = \sum_{j=0}^{\infty} b_j e_{p-j}$ where the $b_j$'s are chosen so that $\hat{\theta}_{p+1}$ is the minimum mean square error estimate of $\theta_{p+1}$. If the same assumptions are made about the measurement errors, then it has already been shown that $\hat{Z}_{p+1}$, the best estimate of the observed value of the quality characteristic in the $(p+1)^{th}$ phase, is equal to $\hat{\theta}_{p+1}$, the best estimate of the optimal setting in the $(p+1)^{th}$ phase. Thus choose the $b_j$'s so that $\hat{Z}_{p+1} = \hat{\theta}_{p+1}$ is the minimum mean square estimate of $Z_{p+1}$.

## The Relationship Between Optimization, Control and Prediction

The prediction problem differs somewhat from the control and optimization problems. In the prediction problem the Z's are observed directly and the predictor $\hat{\theta}_{p+1}$ can be calculated from $\hat{\theta}_{p+1} = \sum_{j=0}^{\infty} a_j Z_{p-j}$ where the a's are chosen to minimize $E(\epsilon_{p+1}^2) = E(\theta_{p+1} - \hat{\theta}_{p+1})^2$. On the other hand, in the optimization and control problems the Z's are not observed directly but rather the e's. Since $Z_p = X_p + e_p$ and since $X_p$

is equal to $(X_{p-1}+x_p) = (X_{p-2}+x_{p-1}+x_p) = \ldots = \sum_{j=0}^{\infty} a_j Z_{p-j-1} = \hat{\theta}_p$, then $Z_p$ can be written as $\hat{\theta}_p + e_p$. Thus the Z's can be reconstructed from the observed deviations if necessary and the predictor $\hat{\theta}_{p+1} = \sum_{j=0}^{\infty} a_j Z_{p-j}$ can be used. Therefore it appears that the three problems are completely analogous. However, the main differences arise in differentiating which quantities are actually observed and which are calculated. In the prediction problem the quantity $Z_p$ is observed and a new quantity $\hat{\theta}_{p+1}$ is calculated which minimizes the squared error between the actual $\theta_{p+1}$ setting and the estimate. In the control problem the actual deviation $e_p$ of the quality characteristic from its target value is observed and a new correction to the process mean $x_{p+1}$ is computed which, when applied, will minimize the deviation from target of this characteristic in the coming phase. Likewise in the optimization problem the slope

$$e_p = \frac{\bar{Y}(X_p+\delta) - \bar{Y}(X_p-\delta)}{2\delta\beta_{11}},$$

which measures the distance of $X_p$ from the optimum, is actually observed. Then an adjustment $x_{p+1}$ to the set-point $X_p$ is computed which minimizes the expected squared deviation of $X_{p+1}$ from its optimal setting.

## A Practical Prediction and Control Model

Box and Jenkins use the above prediction, control and optimization models to develop a formal model which yields optimum solutions to their respective problems under specified conditions. However, their approach is difficult to apply and the conditions under which optimality occur are realistic for only a relatively few time series. Realizing the shortcomings of their formal model, they developed a more practical predictor.

It is well known that a mean projecting predictor, which has proven to be of great value and which has been shown to be related to the prediction and control models already presented, is the exponentially weighted mean. For the predictor model developed earlier, the change in the observed value, $\Delta Z_{p+1} = \hat{Z}_{p+1} - Z_p$, might be written as $\gamma_o e_p$ where $\gamma_o$ is a weighting factor. Box and Jenkins knew from experience with their formal model that for time series involving linear trends the predicted change to be made on the observed value could be written as $\Delta Z_{p+1} = \gamma_o e_p + \gamma_1 S^1 e_p$ where $S^1 e_p = \sum_{j=0}^{p} e_{p-j}$ and $\gamma_1$ is another weighting factor. Therefore, as a natural generalization to the above type of model augmentation, it can be reasoned that for time series involving several different components such as trends or periodic variations the predicted change to be made on the observed value could be written as $\Delta Z_{p+1} = (\gamma_{-L} S^{-L} + \ldots \gamma_{-2} S^{-2} + \gamma_{-1} S^{-1} + \gamma_o + \gamma_1 S^1 + \gamma_2 S^2 + \gamma_m S^m) e_p$ where $S^{-j} e_p = \Delta^j e_p$ and $S^j e_p$ denotes the jth multiple sum over past history of $e_j$.

It is of interest to note that the widely applied predictor obtained by taking an exponentially weighted mean $\hat{Z}_{p+1} = \gamma_o \sum_{j=0}^{\infty} (1-\gamma_o)^j Z_{p-j}$ corresponds to taking the single central term in the above predictor, namely $\gamma_o e_p$. Box and Jenkins note that, keeping in mind the success of the exponential predictor, it might be expected that the simple generalization $\Delta \hat{Z}_{p+1} = \gamma_{-1} \Delta e_p + \gamma_o e_p + \gamma_1 S^1 e_p$ would be adequate for many practical purposes. They observe that experience of two kinds justifies this simplification:

> Our somewhat limited experience in applying this theory to industrial series has shown that for those series so far tried, this generalization has been adequate. In fact so far as prediction is concerned the term in $\Delta e_p$, has not so far been needed. A further vast fund of experience in this area is possessed by control engineers. We have seen already that if there were no dynamics then the adjustment $x_{p+1}$ of the control set-point should be made equal to $\Delta Z_{p+1}$ the predicted change. A form of automatic control commonly used in industrial plants in continuous time makes a correction proportional to a linear combination of (i) the first derivative of the current deviation, (ii) the deviation itself, and (iii) the integral of the deviations over all past history. If, therefore, we were using our predictor for control purposes, we would employ a discrete time analogue of what control engineers have been using for years.

The types of continuous control mentioned in the above quotation are called respectively derivative, proportional and integral control. The three types of control in Box and Jenkin's discrete time analogue are called respectively first difference, proportional and cumulative control.

Thus, the practical Box-Jenkins model is as follows. The predictor of the next value of the time series for one period or phase in the future is $\hat{Z}_{p+1} = Z_p + \gamma_{-1}\Delta e_p + \gamma_o e_p + \gamma_1 S^1 e_p$ where $Z_p$ is the actual value of the series in the pth phase; $\gamma_{-1}$ is the parameter of first difference control; $\gamma_0$ is the parameter of proportional control; $\gamma_1$ is the parameter of cumulative control; and $\Delta e_p$, $e_p$, and $S^1$ are the terms defined earlier in this chapter. A major problem encountered in using this model is in the estimation of the values of the three control parameters. A spectral method of estimation is presented by Box and Jenkins. However a simpler method recommended by Box and Jenkins for practical application of the model is to evaluate the error sum of squares $S(\gamma_{-1},\gamma_0,\gamma_1) = e_1^2 + e_2^2 + \dots + e_n^2$ for a grid of values of the three control parameters. Then select that combination of parameter values which yields the minimum value of the error sum of squares. This parameter combination is then used in subsequent time periods to predict or to control the value of some time series. This crude grid search technique will be used in this thesis to form a basis of comparison with which the two evolutionary operation (EVOP) techniques for estimating the control parameters will be compared. These two EVOP techniques will be discussed in the following chapter.

It should be remembered that the two problems of prediction and control are very similar, differing only by which term, actual series value or the deviation, is actually observed. For purposes of this investigation, it will be sufficient to consider only the predictive use of the model; since with proper transformations of the time series values, the analogous control problem can also be studied.

As a final point it is important to note the ranges of the three control parameters to be considered in this study. Box and Jenkins have established the ranges for the parameters which result in stability for the model to be as follows:

$$-1 \leq \gamma_1 \leq +1;$$

$$0 \leq \gamma_0 \leq +2;$$

$$0 \leq \gamma_1 \leq +2.$$

Stability implies that the model will not lead to infinite prediction variances as has been shown to be possible if the above ranges are exceeded. Therefore in both the crude grid search and in the application of the EVOP techniques, the above ranges will not be exceeded.

# CHAPTER III

## EVOLUTIONARY OPERATION PROCEDURES

This chapter will present a description of the two response surface techniques which will be used to improve the Box-Jenkins model. The name "response surface" was coined by Box to designate surfaces which are formed by the "response" of a certain criterion from various combinations of environmental or independent factors. In this thesis the square of the forecast error is used as the criterion and the three control parameters are the environmental factors. The square of the prediction or forecast error was chosen for several reasons. Most advanced forecasting models use this as the criterion to be minimized. For response surface analysis the square of the error always insures values of the response surface greater than or equal to zero. It also places more emphasis on larger errors tending to make the convex surface more pronounced.

The prediction error is defined as

$$E(t+1) = F(t+1) - X(t+1),$$

where F(t+1) is the forecast made at time t for period t + 1, and X(t+1) is the actual time series value at time t + 1. Thus the square of the forecast error is defined as

$$[E(t+1)]^2 = RS(t+1) = [F(t+1)-X(t+1)]^2$$

where RS(t+1) is the value of the response surface at time t + 1. The rationale is to form a response surface based on several combinations of the control parameters. Predictions are made using the Box-Jenkins model at the various parameter combinations and the resulting errors and response surface values are calculated. Once the response surface is approximated by these values, an analysis can then be performed to determine which combination of control parameters should be used to form the new response surface and what values to use in making the forecast for the next time period.

Several response surface methods are used in various kinds of practical problems; however, the methods to be investigated here are known as Evolutionary Operation (EVOP). These techniques imitate the natural evolutionary process described by Box (1) in that they consist of systematically introducing variation in selected independent variables which affect the process, and then in some manner select the best operating conditions. In this way information is produced in a systematic manner and the results are immediately applied. The perturbations introduced through EVOP are aside from normal process variability; and from information gleaned through this variation, EVOP gradually pushes the process toward its optimal operating conditions.

Two EVOP techniques are to be investigated herein. One procedure is just the usual EVOP described by Box (1) and more recently by Box and Draper (2), which utilizes the $2^k$ factorial experimental design. The other technique, described by Spendley (15), utilizes the simplex design.

## Factorial EVOP

It was shown in Chapter II that the Box-Jenkins model is a three-parameter model, the three parameters being first difference, proportional and cumulative. If this model is to be made self-adaptive, it is necessary to alter these three parameters in response to changes in the nature of the time series under investigation. For example, if the model is operating with a given set of control parameters and at some point the time series begins to exhibit a pronounced linear trend, then the accuracy of the model could possibly be improved if the parameters were permitted to take on different values. The $2^3$ factorial experimental design is one method by which this parameter manipulation might be accomplished. The use of the $2^3$ factorial experimental design allows the system to filter out random fluctuations and, therefore, to respond only to actual changes in the basic nature of the series.

The actual operation of the $2^3$ factorial design is as follows. High and low values are formed about each point of the given set of parameter values. This given set of values at which the model is now operated is called the center point of the experimental design. Associated with this center point is a value of the response surface. For example, the center point might have the control parameter values of $\gamma_{-1} = 1/2$, $\gamma_0 = 1/2$ and $\gamma_1 = 1/2$. If high and low values are taken equally about this center point, the eight points of the design matrix D are formed. Each row in D represents a point at the eight vertices of the cube formed about the center point. A one represents a high value and a zero represents a low value.

$$D = \begin{array}{c} \begin{array}{ccc} \gamma_{-1} & \gamma_0 & \gamma_1 \end{array} \\ \begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 0 \\ 1 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 1 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix} \end{array}$$

Figure 1 presents a graphical display of the center point and the factorial design surrounding it. If the prediction model is operated at each of these eight points, then the corresponding response surface values can be obtained.

The effects of varying each parameter can now be determined. If the effect or effects are statistically significant, then it is desired to move the center point, at which the actual prediction will be made for the next time period, in such a way as to decrease the response. In other words, if the positive effect of increasing $\gamma_{-1}$ is statistically significant, it will hopefully decrease the next period's actual prediction error if the value of $\gamma_{-1}$ is decreased. Likewise, if the positive effects of $\gamma_{-1}$, $\gamma_0$ and $\gamma_1$ are all significant, it is desirable to decrease all three coordinates of the next period's center point. The expressions for the effects of the three parameters are:

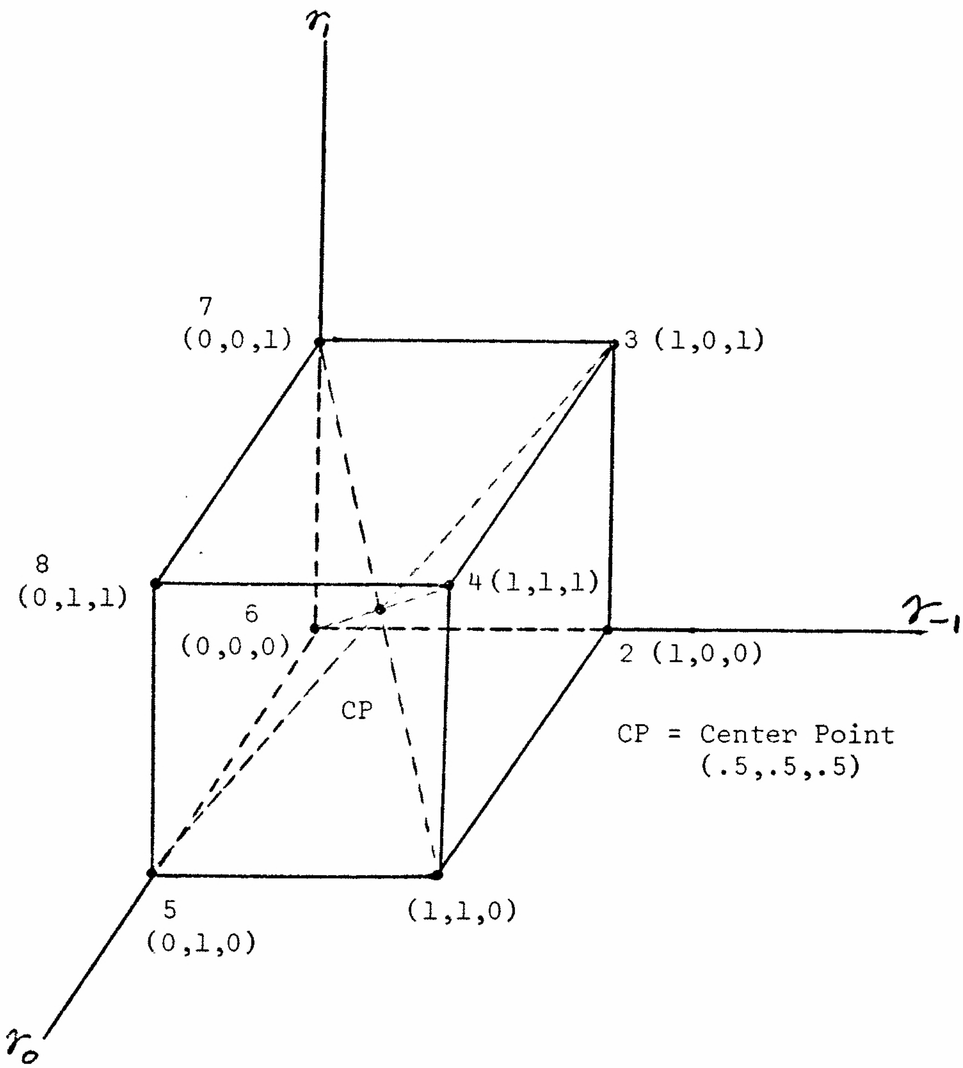


Figure 1. $2^3$ Factorial Experimental Design

$$\text{effect of } \gamma_{-1} = \frac{1}{4}\,[\bar{r}_1+\bar{r}_2+\bar{r}_3+\bar{r}_4-\bar{r}_5-\bar{r}_6-\bar{r}_7-\bar{r}_8];$$

$$\text{effect of } \gamma_{0} = \frac{1}{4}\,[\bar{r}_1+\bar{r}_4+\bar{r}_5+\bar{r}_8-\bar{r}_2-\bar{r}_3-\bar{r}_6-\bar{r}_7];$$

$$\text{effect of } \gamma_{1} = \frac{1}{4}\,[\bar{r}_3+\bar{r}_4+\bar{r}_7+\bar{r}_8-\bar{r}_1-\bar{r}_2-\bar{r}_5-\bar{r}_6]$$

where $\bar{r}_1, \bar{r}_2, \ldots, \bar{r}_8$ are the average response surface values at the respective experimental design points. It will be noticed here that it is necessary only to determine main effects, since no meaningful information would result from an analysis of the interactions.

This $2^3$ factorial design differs somewhat from the usual three-parameter EVOP scheme. Usually a three-variable EVOP involves blocking in order to eliminate time effects ordinarily found in industrial applications. EVOP was mainly developed for the chemical industry and since their experimentation involves changing a process slightly, the data may be affected by time effects in the process. This is especially true where considerable time passes between parameter perturbations. However, in the prediction and control problems, time effects do not exist, since all eight responses can be determined simultaneously.

In this application, 99 per cent confidence intervals on the effects shown above are used instead of the usually recommended 95 per-cent confidence intervals. It was felt that by increasing the size of the confidence interval, the possibility of reacting to chance causes rather than true time series changes was reduced. Thus, in order for a control parameter to be changed, its effect must be significant in a smaller critical region. In this way the control parameters will not tend to fluctuate unnecessarily and the stability of the system is increased. The use of a 99 per cent confidence interval is approximately equivalent to a ±3 standard deviation range. This standard deviation is estimated by a range method presented in the following paragraphs.

## Standard Error of Effects for a $2^3$ Factorial

If n cycles of a $2^p$ design in p factors are performed, each main effect will be a contrast between the average of one half of the observations and the average of the other half. If $\sigma^2$ is the variance of the individual observations, assumed to be independently distributed, then the variance of an average of $\frac{1}{2}(n2^p)$ observations is $\frac{2\sigma^2}{n2^p}$ (see reference 2). Each main effect and interaction is the difference of two such independent differences and, therefore, has variance

$$\frac{2\sigma^2}{n2^p} + \frac{2\sigma^2}{n2^p} = \frac{4\sigma^2}{n2^p} .$$

The standard error of each effect will be obtained by taking the positive square root of the above quantity and substituting S, the sample standard deviation, for the unknown standard deviation $\sigma$. For a $2^3$ factorial design the variance is $\frac{\sigma^2}{2n}$ and the standard error of the effect is $\frac{S}{\sqrt{2n}}$. Thus the value of S, the estimate of $\sigma$, is needed. The method used to estimate $\sigma$ is developed by Box and Draper (2). Let D be the difference of the average and the new observation at any point p in the design; also let $X_{p,i}$ be the observation at point p in cycle i. Let n be the number of cycles. The general expression for the differences used in estimating the standard deviation of response differences at a point is

$$D_p = \frac{X_{p,1} + X_{p,2} + X_{p,3} + \ldots + X_{p,n-1}}{(n-1)} - X_{p,n}.$$

The variance of $D_p$ can be expressed as

$$\sigma^2_{D_p} = \frac{1}{(n-1)^2}\left[S^2_{x,p_1} + S^2_{x,p_2} + \ldots + S^2_{x,p_{n-1}}\right] + S^2_{x,p_n}$$

where $S^2_{x,p_i}$ is the estimate of the variance of each observation at the pth point in the ith cycle. Since the X's come from the same population, the variances are equal and

$$\sigma^2_D = \frac{1}{(n-1)^2}\left[S^2_x(n-1)\right] + S^2_x,$$

or

$$\sigma^2_D = \frac{n}{(n-1)} S^2_x.$$

Thus the standard deviation can be written as $\sigma_D = \sqrt{n/n-1}\, S_x$. In terms of the standard deviation of the differences, the standard deviation of the population is $S_x = \sqrt{n-1/n}\, \sigma_D$. The standard deviation $\sigma_D$ is determined using a range technique. Box and Hunter (3) justify the use of ranges as estimates of $\sigma_D$ because "it is known to be very little less efficient and somewhat more robust than the estimate based on the sum of squares."

Now $\sigma_D$ has been determined in the field of quality control as $\sigma_D = \frac{R_d}{d_2}$ where $R_d$ is the range of differences and $d_2$ is a control chart constant. From Burr (8) $d_2$ is the constant which produces an unbiased estimate of $\sigma_D$ from $R_d$. The values of $d_2$ depend on the number of experimental points in the design. The value of $d_2$ which is used with nine experimental points is 2.970. Hence for a given number of cycles, n, $S_x$ can be written $S_x = \sqrt{n-1/n}\,\frac{R_d}{d_2}$. The quantity $\sqrt{n-1/n}\,\frac{1}{d_2}$ is usually denoted $f_{k,n}$ where k is the number of experimental points. A table of $f_{k,n}$ values can be found in Box and Draper. Successive averages of $S_x$ are formed by accumulating the sum of $S_x$ and dividing by n-1. In this manner the estimate of $S_x$ is continually updated. Thus the three standard deviation limits are $\pm 3S_x/\sqrt{2n}$.

## Sequential Simplex EVOP

An alternative method of Evolutionary Operation is the sequential Simplex technique. This technique was first proposed by Spendley (15), and also discussed by Box and Draper (2). An application of the simplex EVOP technique to sales forecasting is given by Montgomery (11). The sequential simplex technique is simpler than the factorial EVOP technique both conceptually and computationally. However, it has some characteristics which may tend to render it too sensitive to random noise; and hence, its usefulness with the Box-Jenkins model may be limited. One of the objectives of this investigation is to study this possibility.

A simplex is a regularly sided figure consisting of k = n + 1 points in n dimensional space. Thus for n=2 an equilateral triangle is a simplex. In three dimensions a regular tetrahedron is a simplex.

An experimental design in which the design points are the vertices of a simplex is called a simplex design. Like the factorial design a simplex design is an orthogonal, first-order experimental design. The basic idea of the sequential simplex technique can be understood by considering the case for just two variables, say $\gamma_1$ and $\gamma_0$. Assume that it is desired to minimize some quantity such as forecast error which is a function of $\gamma_1$ and $\gamma_0$. The points labeled 1, 2 and 3 on Figure 2 are arranged in

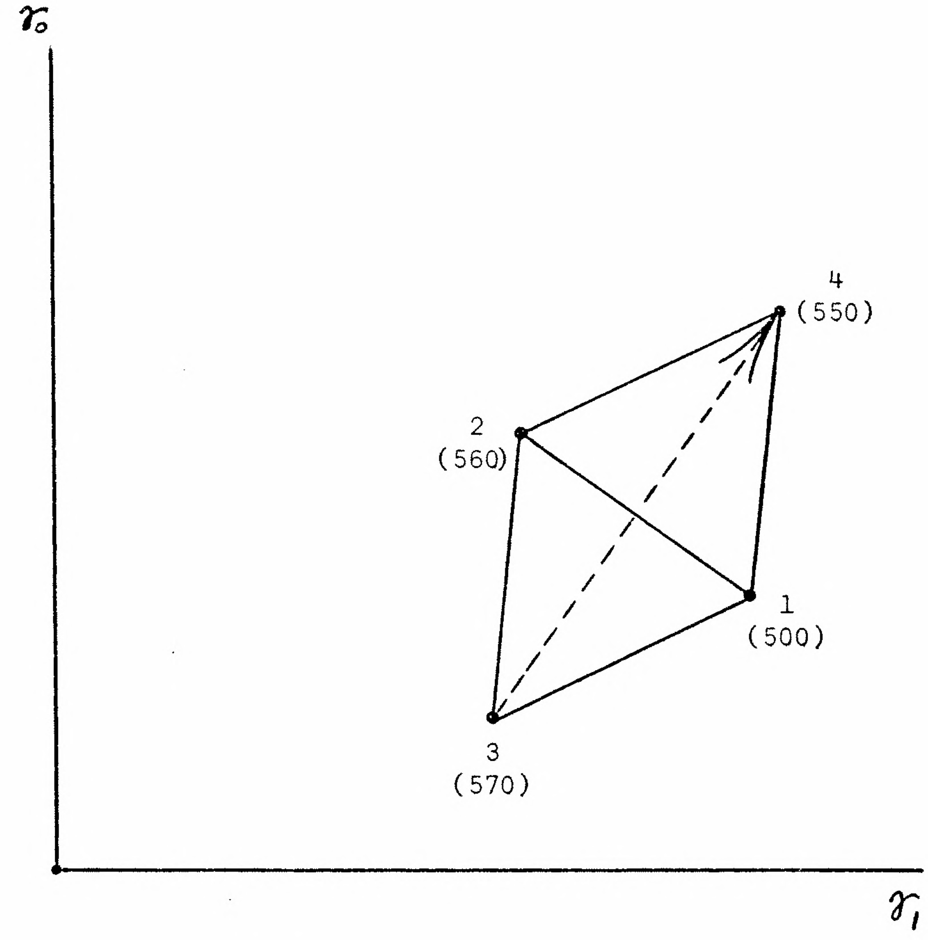


Figure 2. Two Parameter Simplex Design

the form of an equilateral triangle. The values of forecast error for these three runs of the forecasting model are, say, 500, 560 and 570, respectively. Since 570 is the greatest amount of error, the simplex procedure says to delete the point numbered 3 and to locate a new point numbered 4 which forms a new equilateral triangle differing from the previous one by only one point. A run of the forecasting model is now made at the new point yielding an error value of 550. The largest error value now occurs at point 2; so the above procedure is repeated by deleting point 2 and so on. The above procedure is in general the basic operating procedure of the sequential simplex technique. There are some additional rules, however, which will be covered later in this chapter.

To apply the sequential simplex procedure to the Box-Jenkins model it is necessary to use an extension of the above two variable example to three variables. The three variables in this case are the three control parameters in the Box-Jenkins model. In a three parameter problem the number of experimental points needed is four. The experimental arrangement of these four points is in the form of a regular tetrahedron.

The basic design employed is the regular simplex in k dimensions where k is the number of factors or variables under investigation. Relative to a chosen origin $X_1, X_2, \ldots, X_k$, a regular simplex of edge length L is conveniently specified by the (k+1) k design matrix D:

$$\underline{d}'_{j*} = \frac{2}{k}(\underline{d}'_1+\underline{d}'_2+\ldots+\underline{d}'_{j-1}+\underline{d}'_{j+1}+\ldots+\underline{d}'_{k+1}) - \underline{d}'_j.$$

Calculate the prediction for the next period using the control parameters which are the elements of $\underline{d}'_{j*}$.

2. Apply rule "1" unless a design point has occurred in k+1 successive simplexes without being eliminated. Should this situation arise for the ith design point, discard $E_i$ and calculate the prediction for the next period using the control parameters in $\underline{d}'_j$. Then apply rule "1."

3. If $E_i$ is the maximum forecast error squared in the nth simplex and $E_{i*}$ is the maximum error squared in the $(n+1)^{st}$ simplex, then do not return to the original simplex as indicated, since this would result in an immediate reflection to the preceding simplex. Instead, move from the $(n+1)^{st}$ simplex by discarding the second largest value of $E_i$.

It is easily seen that application of the above rules results in a shift in the values of the control parameters at each period. This characteristic may tend to make the design too sensitive to random fluctuations in the time series and thereby lead to an unstable, inaccurate prediction and control system. On the other hand, the factorial design dictates a parameter change only when a statistically significant need to change is shown. Thus some filtering of the noise occurs with the factorial design. However, in the presence of small amounts of random noise, there may be little difference in the stabilities and accuracies of the two systems.

As a closing comment on the two EVOP techniques, it should be noted that the factorial design allows for both technical and empirical feedback while the simplex design allows only empirical feedback. However, since adaptive control and prediction systems are essentially empirical in nature, this deficiency of the simplex design may be no real disadvantage.

# CHAPTER IV

## TEST CRITERIA AND TIME SERIES DESCRIPTION

To compare the basic Box-Jenkins model with its two EVOP modifications, it is necessary to establish a criterion by which to judge the merits of each system. The criterion selected is the accuracy of each system. Accuracy is perhaps the most important measure of any prediction and control system because it determines to a large extent the success or failure of the system. In this thesis accuracy is measured by three quantities: (1) the average error, (2) the corrected sum of squared error, and (3) the sample error variance.

### The Average Error

The average error assesses the bias of the prediction technique. If a system is leading or lagging the actual time series, it will be indicated through the average error. The average error is defined as

$$\bar{E} = \frac{\sum_{t=1}^{k} (F_t - X_i)}{k}$$

where $\bar{E}$ is the average forecast error, $X_t$ is the actual series value at time t, $F_t$ is the forecast value made for time t, and k is the number in the number of series values under consideration. If $\bar{E}$ is greater than zero, then the forecasting system tends to lead the actual time series values. On the other hand, if $\bar{E}$ is less than zero, the forecasting system tends to lag behind the actual time series values. Clearly it is desirable to have the value of $\bar{E}$ as close as is possible to zero, since this would indicate that on the average the prediction system neither leads nor lags behind the actual time series.

## The Corrected Sum of Squared Errors

The average error supplies a measure of the bias of the prediction system; however, another measure is needed which gives an absolute measure of the amount of error in the system. The measure chosen to supply this information is the corrected sum of squared forecast errors. This quantity is defined as CSSE = $\sum_{t=1}^{k}(F_t - X_t - \bar{E})^2$ where CSSE is the corrected sum of squared forecast errors; $F_t$ is the forecast made for time t; $X_t$ is the actual time series value at time t; $\bar{E}$ is the average error; and k is the number of actual series values used. It can be shown that for computational purposes the corrected sum of squared errors can be expressed as

$$\text{CSSE} = \left[ \sum_{t=1}^{k}(F_t - X_t)^2 - \frac{\left(\sum_{t=1}^{k}(F_t - X_t)\right)^2}{k} \right].$$

It is desirable to have a small value of CSSE since this indicates a small amount of error in the prediction technique.

## The Sample Error Variance

Another measure of the amount of variability in the errors resulting from the prediction system is the sample error variance. A prediction system may on the average possess little bias; however, its consistency to predict may be faulty. This is due to the fact that when computing the average error, positive and negative errors tend to cancel. The model should accurately represent the data and the deviation of the forecast from the true series should be small. The sample error variance is the most popular measure of forecast accuracy found in the literature and most authors attempt to formulate their prediction models to minimize this measure. The sample error variance is computed as

$$\mathrm{EVAR} = \frac{\sum_{t=1}^{k} (F_t - X_t - \bar{E})^2}{k - 1}$$

where the terms are those defined previously. It is of interest to note that the sample error variance can be easily computed by dividing the CSSE by one less than the number of time periods in the series.

In a succeeding chapter are presented the results of an analysis in which the above statistical measures of accuracy are compared. From this analysis it is possible to draw conclusions regarding which of the forms of the prediction model is the most accurate. In this analysis the statistical measures are computed for all three forms of the model using seven time series. These seven time series are generated on the computer and possess characteristics of series found in industrial applications. The description of these seven time series is given later in this chapter.

Since the three forms of the Box-Jenkins model involve different methods of setting and updating the three control parameters, it may be helpful to review which parameters are actually used in making predictions from which the statistical measures are computed. In the basic Box-Jenkins model the control parameters are established only once by a crude grid search technique and never altered afterwards. Thus the statistical measures are computed from the errors incurred by operating the model using these initial parameters. However, in the factorial and simplex EVOP techniques, the control parameters change frequently. In the factorial EVOP procedure the parameter values used to forecast are those defining the center point of the factorial experimental design. When the procedure indicates the need for a parameter change, it is the location of the center point that changes and a new experimental design is placed about it. In the simplex EVOP procedure the actual forecast each period is made using that parameter combination existing at the new point which is added to the simplex each period.

## Description of Time Series Used

In this section a description of the time series used in this investigation is presented. As was stated earlier these time series contain the actual values of the quantity to be predicted. The series were generated on the Rich Electronic Computer Center's UNIVAC 1108 computer. Artificially generated time series were used because in these series it was possible to build in specific characteristics. If the characteristics are known, then it is possible to draw general conclusions about the performance of the prediction models on series possessing these characteristics. Conversely, if these characteristics are unknown only restricted conclusions about the particular series at hand can be made.

In this study seven time series were generated, each possessing certain characteristics. Each of these series has a basic deterministic form, but random variation has been superimposed to make the series more representative of actual industrial series. The random noise was supplied by the RANDN library program contained on the UNIVAC 1108 computer. RANDN is a program which generates normally distributed random variables with the distribution parameters specified by the user. Only normally distributed random variables were used in this thesis purely for the sake of simplicity. It is hoped that the conclusions drawn from this study will be equally as applicable to series exhibiting other forms of random variation.

The following paragraphs present a description of the seven time series to include their characteristics and series values. In the remainder of this thesis these series will be referred to by number only. Each of these time series was arbitrarily set as being 100 time units in length.

Series 1

This series is basically a constant one with random noise superimposed on it. The form of the generator equation is as follows:

$$SV(t) = C + RV$$

where SV(t) is the series value at time t, C is the constant component, and RV is the random component. In this series the value of C is 100 and the random component is distributed normally with a mean of 0.0 and a standard deviation of 5.0. Figure 3 shows a portion of this time series.

Series 2

Series 2 is essentially the same as series 1 except that there is a step increase in the constant component at time period 50. The form of the generator equation is as follows:

$$SV_1(t) = C_1 + RV \quad \text{for } t \leqq 50;$$

$$SV_2(t) = C_2 + RV \quad \text{for } t = 51 \text{ to } 100;$$

where $SV_1(t)$ and $SV_2(t)$ are the time series values for time periods 1 to 50 and 51 to 100, $C_1$ and $C_2$ are the constant components for the two time intervals, and RV is the random component. In this series $C_1 = 50$, $C_2 = 100$ and the random component is normally distributed with a mean of 0.0 and a standard deviation of 5.0. Figure 4 shows a portion of this time series.

Series 3

This series is identical to series 1 except that there is an impulse at time 100. The constant component and the random component are identical to those of series 1. At time 50, however, the constant component has a value of 200 instead of 100. Figure 5 shows a portion of this time series.

Series 4

This series contains a constant component and a linear trend component plus the superimposed random noise. The form of the generator equation is as follows:

$$SV(t) = C + Bt + RV$$

where SV(t) is the series value at time t; C is the constant component; B is the trend per time period; and RV is the random component. For this series C has a value of 10, B has a value of 5, and the random component is normally distributed with a mean of 0.0 and a standard deviation of 3.0. Figure 6 shows a portion of this time series.

Series 5

Series 5 was intended to be a series exhibiting strong seasonal or periodic variation. To generate such a time series the following equation was used:

$$SV(t) = A + B\text{Sin}\,\frac{(2\pi t)}{T} + RV$$

where SV(t) is the series value at time t, A is the constant component, B is the amplitude of the sinusoid, T is the period of the sinusoid, and RV is the value of the random component. In this series A has a value of 50, B has a value of 20, T a value of 12, and RV is normally distributed with a mean of 0.0 and a standard deviation of 6.0. Figure 7 shows a portion of this time series.

Series 6

Series 6 is composed of constant, trend and periodic components. The generator equation is as follows:

$$SV(t) = A + Bt + C\text{Sin}\,\frac{(2\pi t)}{T} + RV$$

where SV(t) is the series value at time t, A is the constant component, B is the linear trend component per time period, C is the amplitude of the sinusoid, T is the period of the sinusoid, and RV is the random component. For this series A has a value of 10, B has a value of 5, C a value of 20, T a value of 12 and RV is normally distributed with a mean of 0.0 and a standard deviation of 5.5. Figure 8 shows a portion of this time series.

Series 7

This series is one that exhibits a high degree of autocorrelation. The form of the generator equation is as follows:

$$SV(t) = \sum_{I=0}^{11} ARV(t+I)$$

where SV(t) is the series value at time t, and ARV(t+I) is an array of random numbers which are normally distributed with a mean of 10.0 and a standard deviation of 6.0. In other words, each value of the time series is found by summing 12 random variables. However, each value of the series differs from the immediately preceding one by only one random component. For example:

$$SV(1) = ARV(1) + ARV(2) + \dots + ARV(12)$$

$$SV(2) = ARV(2) + ARV(3) + \dots + ARV(13).$$

Figure 9 shows a portion of this autocorrelated time series.

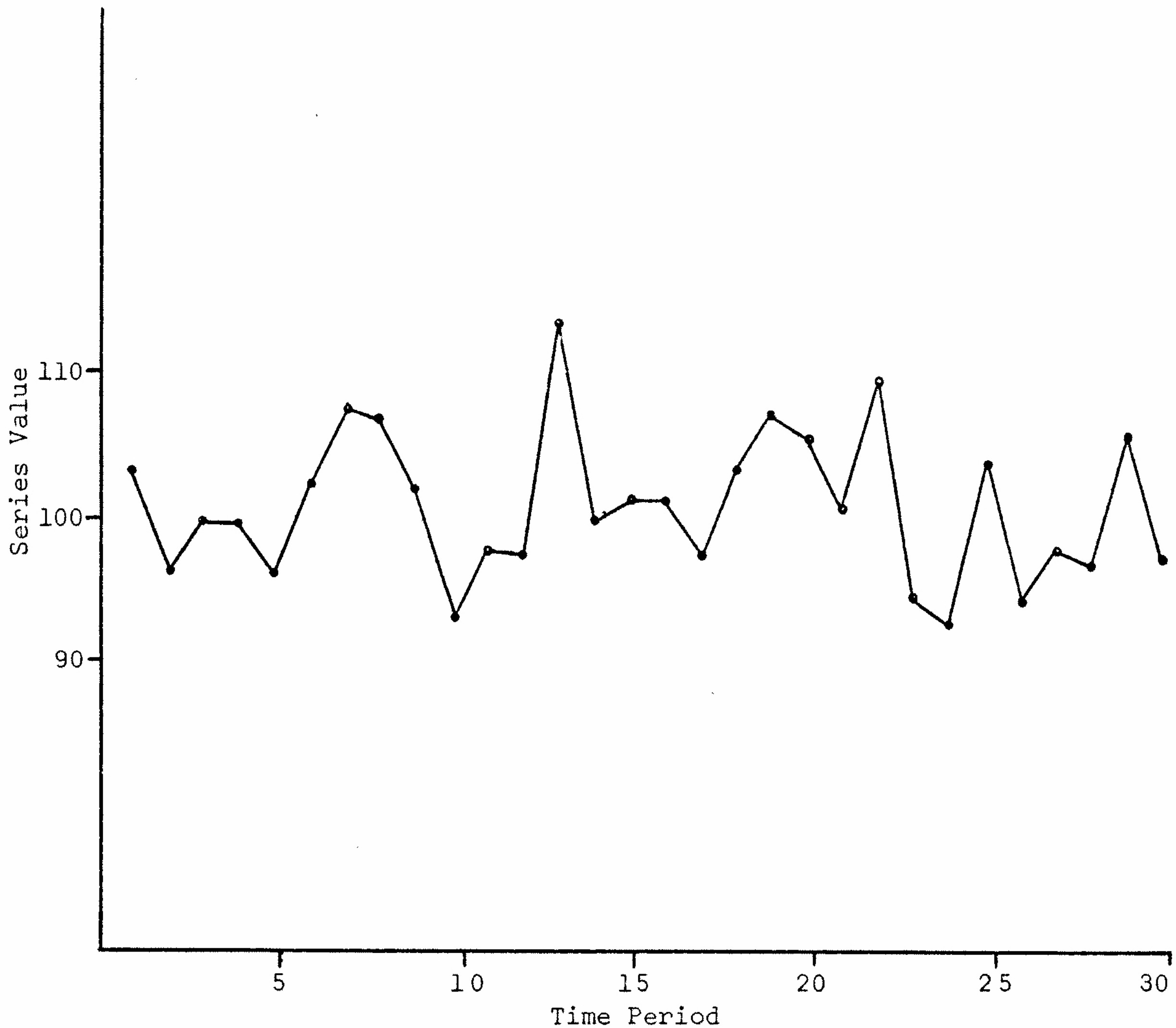


Figure 3. Time Series One--Constant with Random Variation

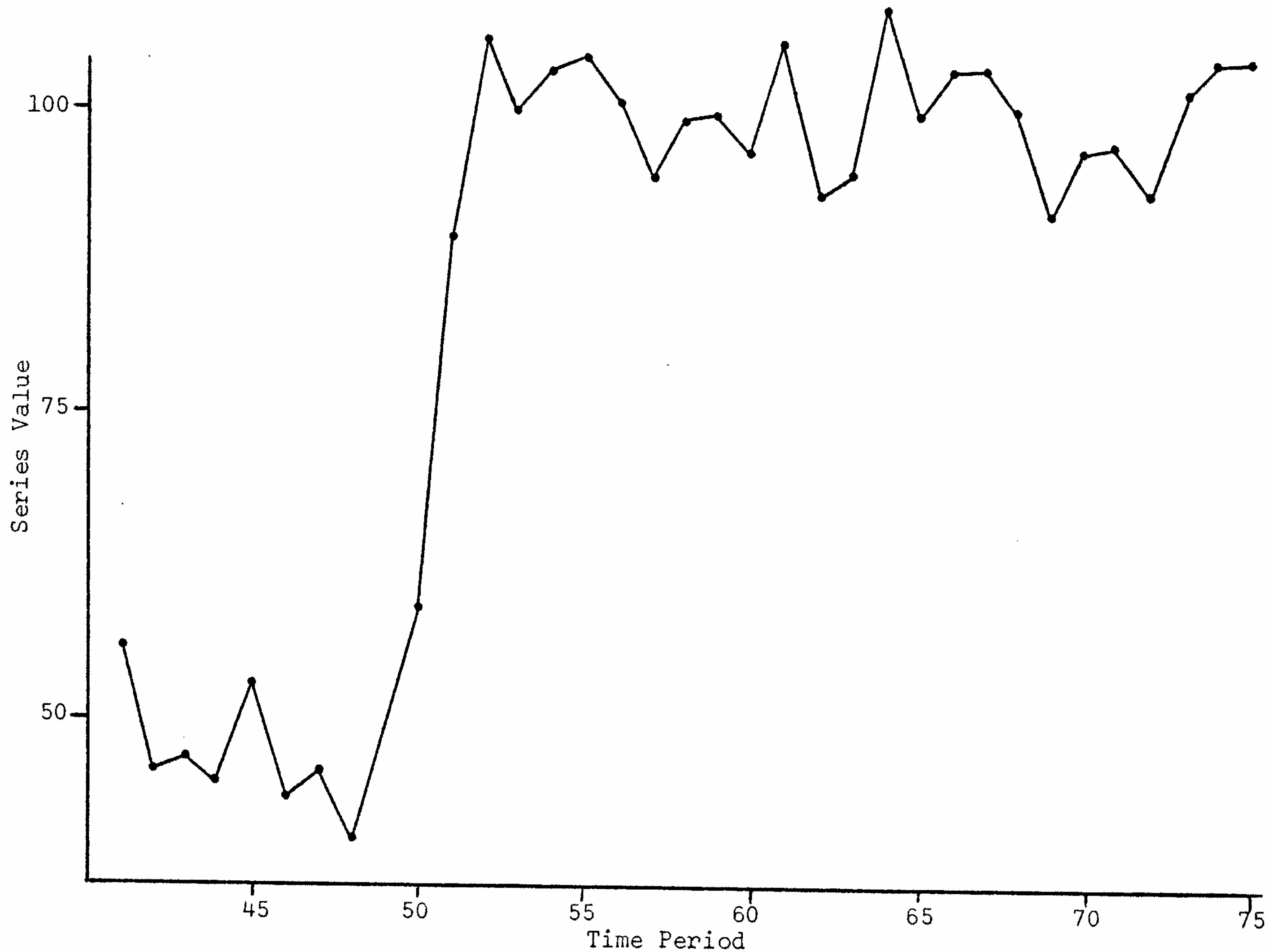


Figure 4. Time Series Two—Step Increase with Random Variation

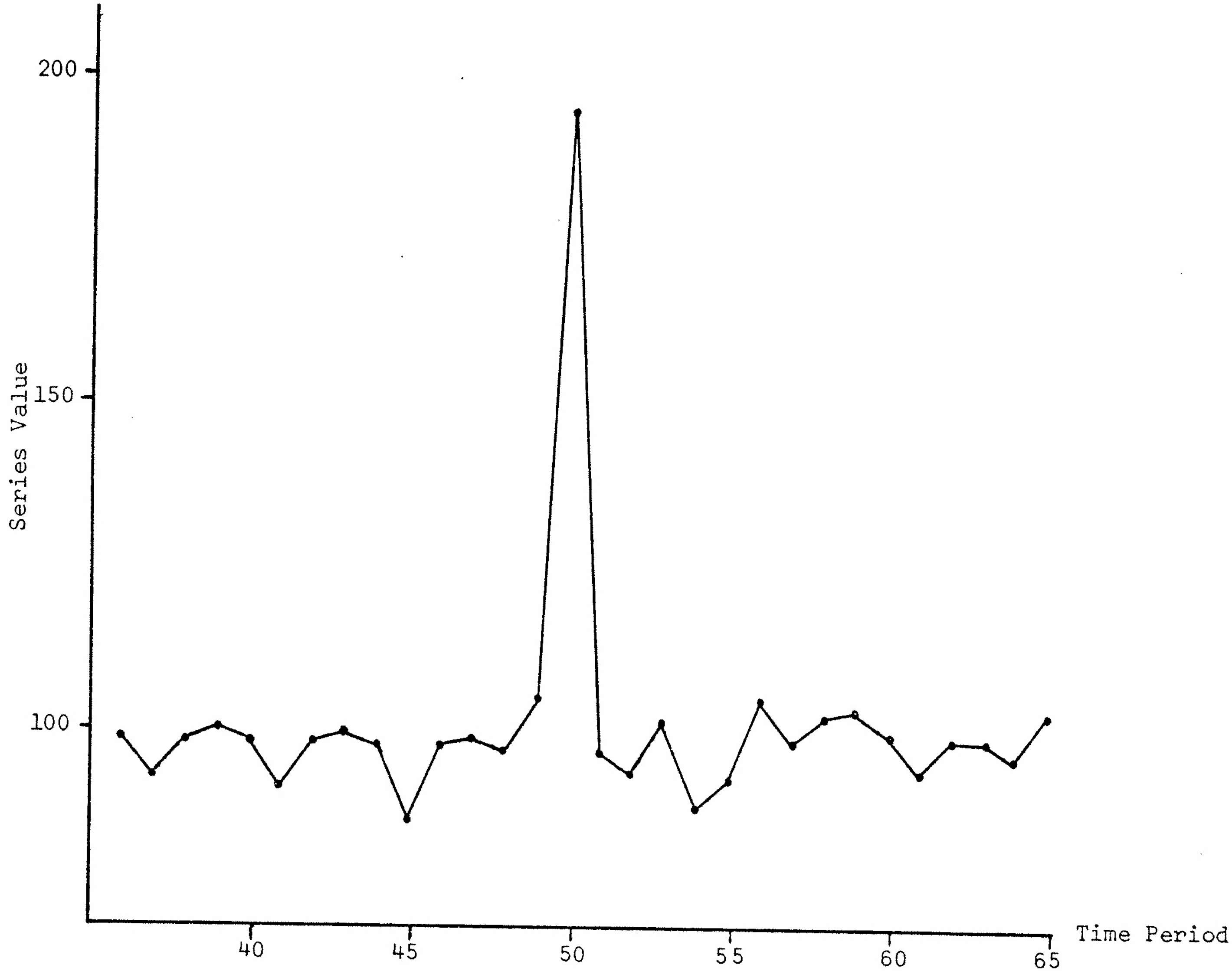


Figure 5. Time Series Three--Impulse with Random Variation

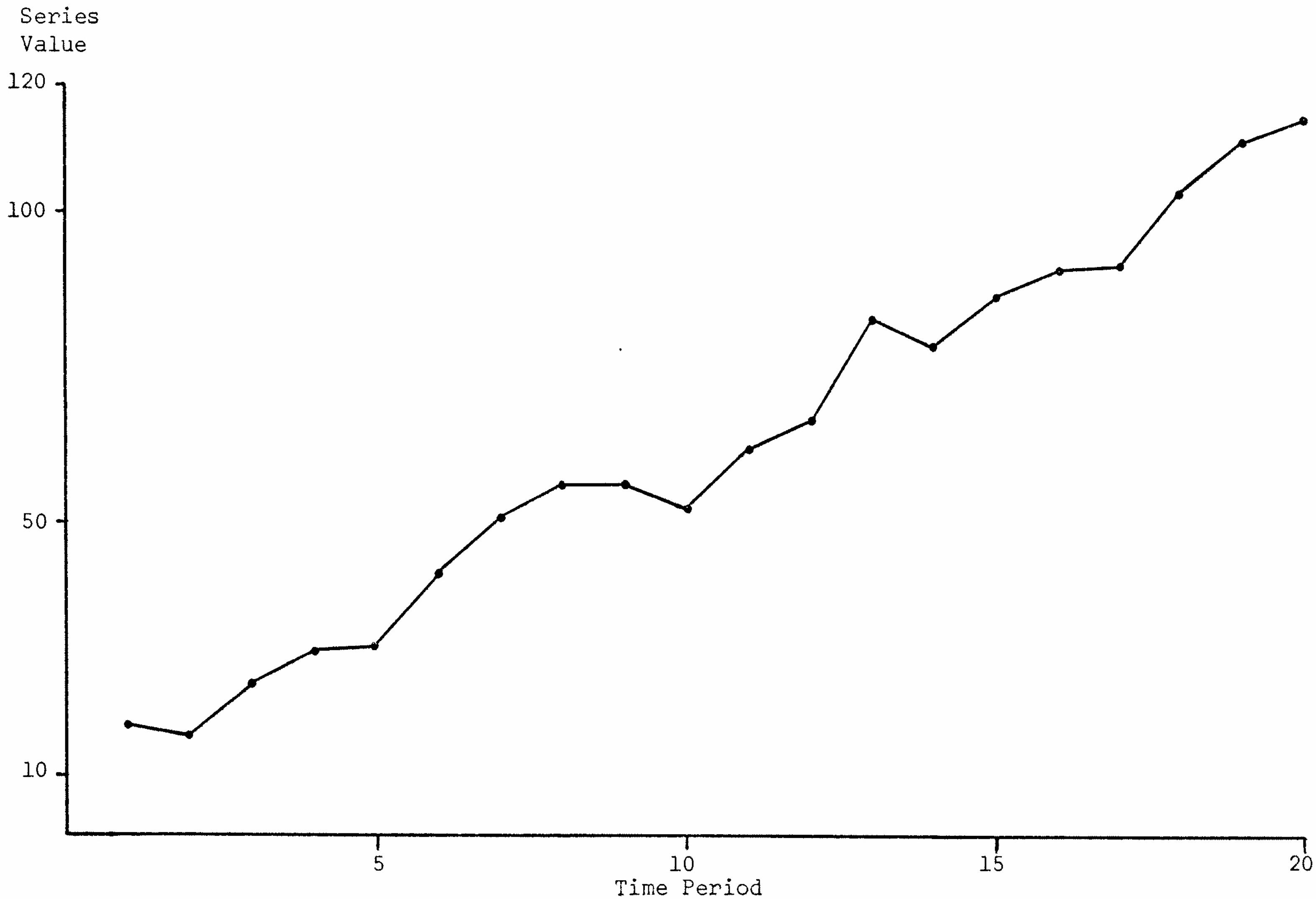


Figure 6. Time Series Four--Linear Trend with Random Components

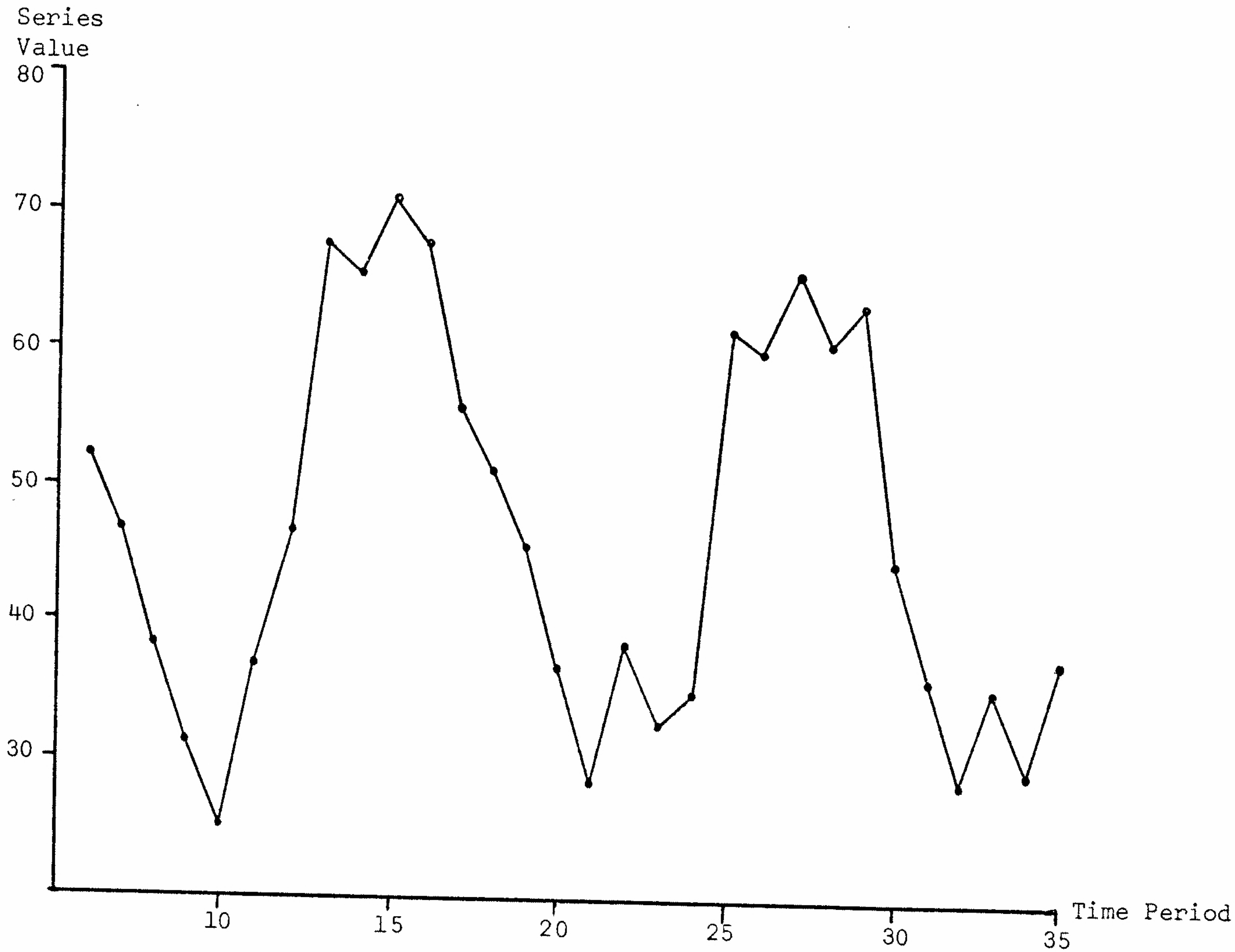


Figure 7. Time Series Five--Sinusoid with Random Variation

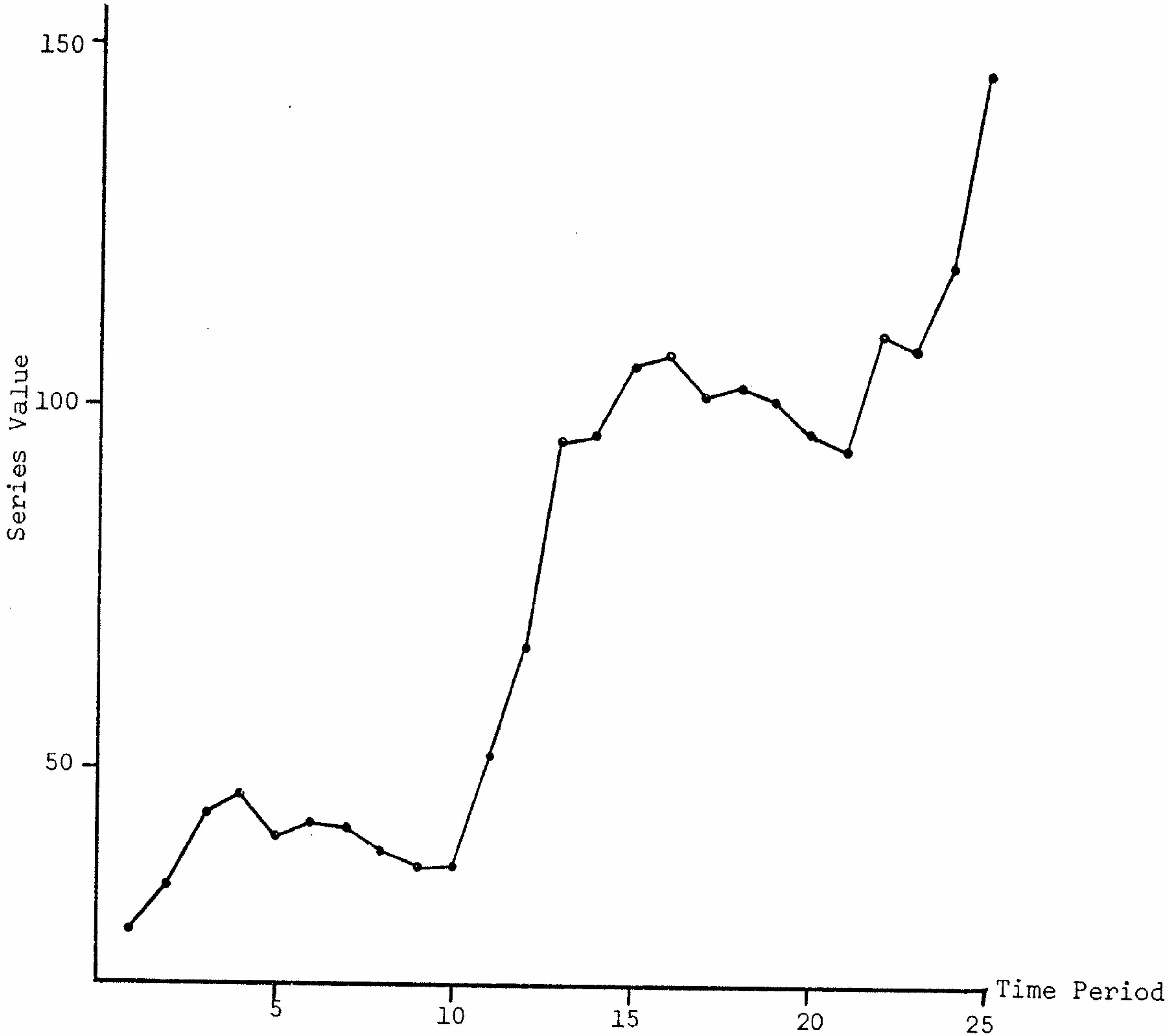


Figure 8. Time Series Six--Linear Trend Plus Sinusoid with Random Variation

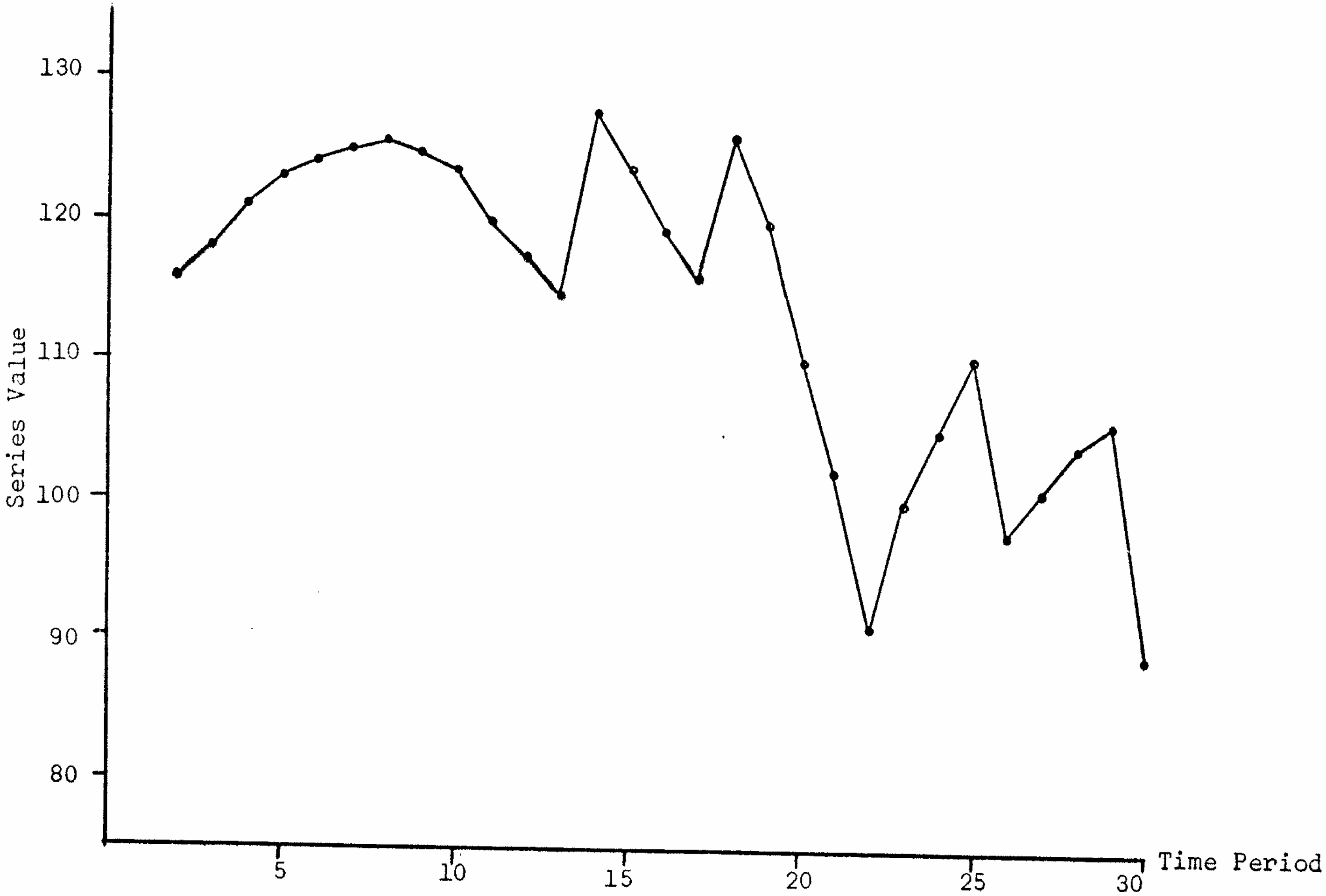


Figure 9. Time Series Seven--Highly Autocorrelated Series

# CHAPTER V

## PRESENTATION OF RESULTS

This chapter presents the results obtained by applying the basic Box-Jenkins model and its two EVOP modifications in predicting the seven time series described in the previous chapter. The measures of prediction accuracy will be discussed first. Then a description of the statistical procedures used to compare the three models will be given along with the results of this comparison. Finally, conclusions will be drawn regarding the relative accuracies of the three prediction procedures.

### Statistical Measures

The three procedures were used to predict each of the seven time series described in the preceding chapter. The errors resulting from these predictions were used to compute the statistical measures of accuracy described in Chapter IV. These statistical measures are summarized in Tables 1 through 3.

### Operational Procedures

A few general comments regarding operational procedures used in the running of the forecasting programs will be helpful. A crude grid search procedure was used with the basic Box-Jenkins model to establish the best values of the control parameters. The first 15 time periods from each time series were used in this crude grid search and the parameter values were selected which yielded the minimum sum of squared errors

Table 1. Average Prediction Error

| Time Series | Basic Box-Jenkins Model | Box-Jenkins Model with Factorial EVOP | Box-Jenkins Model with Simplex EVOP |
|---|---|---|---|
| 1 | .261 | .147 | .208 |
| 2 | -2.329 | -.964 | -2.767 |
| 3 | .390 | 1.997 | .440 |
| 4 | -4.930 | -4.726 | -4.834 |
| 5 | 28,571.520 | -.416 | -89,046.940 |
| 6 | -4.336 | -4.954 | -4.748 |
| 7 | .645 | .643 | .642 |

Table 2. Corrected Sum of Squared Error

| Time Series | Basic Box-Jenkins Model | Box-Jenkins Model with Factorial EVOP | Box-Jenkins Model with Simplex EVOP |
|---|---|---|---|
| 1 | 5,708.918 | 4,392.445 | 5,691.570 |
| 2 | 11,242.390 | 8,259.504 | 11,217.320 |
| 3 | 23,237.500 | 21,927.030 | 23,169.630 |
| 4 | 23,289.250 | 21,089.170 | 20,709.980 |
| 5 | $6{,}867{,}147.000\times10^{6}$ | 17,335.990 | $7{,}812{,}614.700\times10^{6}$ |
| 6 | 793,531.600 | 21,896.210 | 859,649.500 |
| 7 | 6,796.148 | 6,724.260 | 6,804.266 |

Table 3. Sample Error Variance

| Time Series | Basic Box-Jenkins Model | Box-Jenkins Model with Factorial EVOP | Box-Jenkins Model with Simplex EVOP |
|---|---|---|---|
| 1 | 58.254 | 44.821 | 58.077 |
| 2 | 114.718 | 84,281 | 144.462 |
| 3 | 237.117 | 223.745 | 236.425 |
| 4 | 237.645 | 215.195 | 211.326 |
| 5 | $7{,}007{,}286.0\times10^{4}$ | 176.898 | $7{,}972{,}055.81\times10^{4}$ |
| 6 | 8,097.258 | 223.430 | 8,771.934 |
| 7 | 69.348 | 68.622 | 69.431 |

for these 15 time periods. These parameter values were not changed afterwards in the basic Box-Jenkins model. However, the statistical measures include the results of operating the model during the first 15 time periods with these best parameter values. In order that the two EVOP procedures not be penalized, these best parameter values were used as the initial parameter values in both the factorial and simplex schemes. These best parameter values selected by the crude grid search procedure are presented in Table 4 which appears on the following page.

Another important question is the sensitivity of the simplex procedure to the edge length utilized, and the sensitivity of the factorial procedure to the "spread" or upper and lower parameter limits of the experimental design. To determine the proper edge length for the simplex procedure, runs were made using edge lengths of 0.07, 0.03, 0.01, 0.005 and 0.001. Table 5 presents the crude sums of squared errors resulting from these trials.

Table 4. Best Parameters Resulting from Crude Grid Search

| Time Series | Constant of First Difference Control | Constant of Proportional Control | Constant of Cumulative Control |
|---|---|---|---|
| | CONTROL PARAMETERS | | |
| 1 | 0.70 | 0.90 | 0.00 |
| 2 | 0.70 | 0.80 | 0.00 |
| 3 | 0.70 | 0.90 | 0.00 |
| 4 | 0.40 | 0.00 | 0.00 |
| 5 | 0.90 | 0.20 | 0.90 |
| 6 | 0.50 | 0.00 | 0.00 |
| 7 | 0.30 | 0.60 | 0.00 |

Table 5. Crude Sum of Squared Errors for the Seven Time Series Using Various Edge Lengths of the Simplex Design

| Time Series | EDGE LENGTH .07 | .03 | .01 | .005 | .001 |
|---|---|---|---|---|---|
| 1 | 6,945.12 | $2.07\times10^{12}$ | $7.86\times10^{9}$ | 5,946.7 | 5,748.60 |
| 2 | $1.017\times10^{10}$ | $3.11\times10^{6}$ | 26,560.63 | 16,058.55 | 11,223.72 |
| 3 | $1.006\times10^{9}$ | $3.62\times10^{5}$ | 23,265.11 | 23,407.90 | 23,198.07 |
| 4 | $1.18\times10^{6}$ | 42,448.8 | 20,902.64 | 20,784.68 | 20,734.12 |
| 5 | $2.54\times10^{19}$ | $2.40\times10^{14}$ | $2.20\times10^{13}$ | $1.25\times10^{13}$ | $7.81\times10^{12}$ |
| 6 | $9.08\times10^{10}$ | $2.72\times10^{9}$ | $2.33\times10^{6}$ | $4.22\times10^{6}$ | 859,675.62 |
| 7 | $2.15\times10^{6}$ | $6.06\times10^{5}$ | 6,924.69 | 6,846.17 | 6,809.27 |

It was decided that the system is relatively sensitive to the edge length, and the edge length used to calculate the statistics in Tables 1 through 3 was .001. A similar investigation into the spread of the factorial experimental design was made using spreads of 0.50, 0.10, 0.07 and 0.05. Here *spread* is defined as the distance from the center point of the design to the face of the cube formed by the eight experimental points. The results of these runs are presented in Table 6.

Table 6. Crude Sum of Squared Errors for the Seven Time Series Using Various Spreads Values of the Factorial Design

| Time Series | SPREAD .50 | .10 | .07 | .05 |
|---|---|---|---|---|
| 1 | 5,201.72 | 6,749.32 | 4,410.22 | 4,394.93 |
| 2 | 8,429.23 | 10,127.94 | 9,121.29 | 8,351.77 |
| 3 | 22,943.72 | 26,492.91 | 24,193.96 | 22,021.25 |
| 4 | 23,969.74 | 24,289.82 | 29,243.71 | 23,302.18 |
| 5 | 19,442.48 | 19,998.65 | 18,293.52 | 17,409.26 |
| 6 | $7.98\times10^{8}$ | $3.24\times10^{4}$ | 23,249.38 | 23,144.82 |
| 7 | 11,524.86 | 10,493.18 | 7,958.98 | 6,942.53 |

They indicate that a spread of 0.05 is appropriate. These values of edge length and spread were chosen arbitrarily for testing and are in no way meant to be optimal. They were chosen purely to give some relative indication of the sensitivity of the models to these variables. In actual practice the best values of the edge length or the spread might be established either by simulation and analysis of historical data or by trial and error experimentation in real time. Clearly the former method of setting these values would be the most desirable in an operating industrial prediction or control system.

## Evaluation of the Data

In order to evaluate the statistical measures to determine the relative accuracies of the three prediction techniques, a series of statistical tests were performed. A non-parametric or distribution-free approach was used because it was felt that a major assumption in normal theory did not apply. This assumption is that the populations do not have the same variance. Siegel (16) points out that this condition is one which must apply before parametric tests can be used successfully.

The data were analyzed using the Wilcoxon signed rank test. The Wilcoxon test is "a test for the median of a single sample in which we give ranks to the absolute magnitude of the observations and then give to the ranks the signs of the corresponding observations" (6). Brownlee (6) explains that the test is based on the total number of ways ranked sums can be produced. For a test with n differences, the total number of subsets is $2^n$ if any rank is equally likely. Thus in the situation involving small n, the critical region of the test can be developed by complete enumeration of the possibilities. For a one-sided test we would want all the ranked sums less than or equal to some value. Ostle (12) presents the following step-by-step procedure which is used in this thesis.

1. Rank the differences without regard to sign, that is, rank the absolute values of the differences. (The smallest difference is given rank 1 and ties are assigned average ranks.)

2. Assign to each rank the sign of the observed difference.

3. Obtain the sum of the negative ranks and the sum of the positive ranks.

4. Denote by T the absolute value of the smaller of the two sums of ranks found in the previous step.

5. To test the hypothesis of no difference between the effects of the two treatments, compare T with the tabulated critical values.

6. If the observed value of T is less than or equal to the tabulated value, the hypothesis is rejected; otherwise, it is not rejected.

The Wilcoxon signed rank test is used to compare the relative measures of accuracy of the three prediction procedures. First the basic Box-Jenkins model is compared with its factorial EVOP modification. Next the basic Box-Jenkins model is compared with its simplex EVOP modification. Finally the factorial and simplex EVOP modifications are compared with each other. These comparisons are performed using only the average error and the sample error variance measures of accuracy. A comparison of the corrected sums of squared errors yields no additional information since this measure is only a multiple of the sample error variance.

<u>Comparison of the Basic Box-Jenkins Model with the Factorial EVOP Modification</u>

The data in Table 7 contains the information relative to this test. The hypothesis to be tested is that the basic Box-Jenkins model is no worse than the factorial EVOP modification. Hence the alternative hypothesis is that the factorial EVOP modification is better than the basic Box-Jenkins model. Here *better* means: (1) has forecast error closer to zero, and (2) has a smaller sample variance. The level of significance chosen for this test was 0.025 and N, the number of matched

Table 7. Basic Box-Jenkins Model Versus Box-Jenkins Model with Factorial EVOP [*Hypothesis:* Basic Model is no worse than Factorial EVOP Modification. N = 7; $\alpha$ = .025; $T_c$ = 2.]

| Time Series | Basic Box-Jenkins Model | Box-Jenkins Model with Factorial EVOP | Difference | Absolute Rank | Positive Rank | Negative Rank |
|---|---|---|---|---|---|---|
| *TEST A. Average Prediction Error* | | | | | | |
| 1 | .261 | .147 | .114 | 2 | 2 | - |
| 2 | -2.329 | -.964 | 1.365 | 5 | 5 | - |
| 3 | .390 | 1.997 | -1.607 | 6 | - | 6 |
| 4 | -4.930 | -4.726 | .204 | 3 | 3 | - |
| 5 | 28,571.520 | -.416 | 28,571.104 | 7 | 7 | - |
| 6 | -4.336 | -4.954 | -.618 | 4 | - | 4 |
| 7 | .645 | .643 | .002 | 1 | 1 | - |
| | | | | | | $T = 10 > T_c = 2$ Accept Hypothesis |
| *TEST B. Sample Error Variance* | | | | | | |
| 1 | 58.254 | 44.821 | 13.433 | 3 | 3 | - |
| 2 | 114.718 | 84.281 | 30.437 | 5 | 5 | - |
| 3 | 237.117 | 223.745 | 13.372 | 2 | 2 | - |
| 4 | 237.645 | 215.195 | 22.450 | 4 | 4 | - |
| 5 | $7{,}007{,}286.0\times10^4$ | 176,898 | $7{,}007{,}109.100\times10^4$ | 7 | 7 | - |
| 6 | 8,097.258 | 223.430 | 7,873.828 | 6 | 6 | - |
| 7 | 69.348 | 68.622 | .726 | 1 | 1 | - |
| | | | | | | $T = 0 < T_c = 2$ Reject Hypothesis |

observations, is 7. From the Wilcoxon table found in Ostle (12), which gives critical values of the sampling distribution for the sum of ranks, the critical value of T is $T_c = 2$. Since this is a one-tailed test, the critical region is those values of T which are so small that the probability associated with their occurrence under the null hypothesis is less than or equal to 0.025.

The differences contained in Table 7 are obtained by subtracting the absolute value of the relevant statistics of the Box-Jenkins model with factorial EVOP from those of the basic Box-Jenkins model. For the case involving the average forecast error, one cannot reject the null hypothesis since T = 10 is greater than the critical $T_c$ value of 2. However, for the case of the sample error variance the null hypothesis can be rejected since T = 0 which is less than the critical $T_c$ value of 2. Thus it may be concluded that the Box-Jenkins with the factorial EVOP is better than the basic Box-Jenkins model in terms of the sample error variance.

### Comparison of the Basic Box-Jenkins Model with the Simplex EVOP Modification

Again the null hypothesis to be tested here is that the basic Box-Jenkins model is no worse than its simplex EVOP modification. Again N has a value of 7 and the level of significance is 0.025. Thus $T_c$, the critical value of T, is equal to 2. Table 8 contains the results of this test. It can be seen that both in the cases of average forecast error and the sample error variance the null hypothesis cannot be rejected since T has a value of 8 for the average error and a value of 14 for the sample error variance, both of which are greater than the

Table 8. Basic Box-Jenkins Model Versus Box-Jenkins Model with Simplex EVOP

[*Hypothesis:* Basic Model is no worse than Simplex EVOP Modification. $N = 7$; $\alpha = .025$; $T_c = 2$.]

| Time Series | Basic Box-Jenkins Model | Box-Jenkins Model with Factorial EVOP | Difference | Absolute Rank | Positive Rank | Negative Rank |
|---|---|---|---|---|---|---|
| *TEST A.* | *Average Prediction Error* | | | | | |
| 1 | .261 | .208 | .053 | 3 | 3 | - |
| 2 | -2.329 | -2.767 | -.438 | 5 | - | 5 |
| 3 | .390 | .440 | -.050 | 2 | - | 2 |
| 4 | -4.930 | -4.834 | .096 | 4 | 4 | - |
| 5 | 28,571.520 | -89,046.940 | -60,475.420 | 7 | - | 7 |
| 6 | -4.336 | -4.748 | -.412 | 5 | - | 5 |
| 7 | .645 | .642 | .003 | 1 | 1 | |
| | | | | | $T=8>T_c=2$ Accept Hypothesis | |
| *TEST B.* | *Sample Error Variance* | | | | | |
| 1 | 58.254 | 58.077 | .177 | 2 | 2 | |
| 2 | 114.718 | 114.462 | .256 | 3 | 3 | |
| 3 | 237.117 | 236.425 | .692 | 4 | 4 | |
| 4 | 237.645 | 211.326 | 26.319 | 5 | 5 | |
| 5 | $7{,}007{,}286.0\times10^4$ | $7{,}972{,}055.81\times10^4$ | $-964{,}769.81\times10^4$ | 7 | - | 7 |
| 6 | 8,097.258 | 8,771.934 | -674.676 | 6 | - | 6 |
| 7 | 69.348 | 69.431 | -.083 | 1 | - | 1 |
| | | | | | | $T = 14 > T_c = 2$ Accept Hypothesis |

critical value of 2. Thus it cannot be concluded that the simplex EVOP modification is better than the basic Box-Jenkins model.

Comparison of the Simplex EVOP Modification with the Factorial EVOP Modification

The null hypothesis to be tested is that the simplex EVOP modification is no worse than the factorial EVOP modification. Again N has a value of 7 and the level of signifiance is .025. Table 9 presents the results of this test. For the average prediction error it is not possible to reject the null hypothesis since T has a value of 10 which is greater than the critical value of 2. However, for the case of the sample error variance it is possible to reject the null hypothesis since T has a value of 2 which is exactly equal to the critical value of 2. Therefore, it can be concluded that the factorial EVOP modification is better than the simplex EVOP modification in terms of the sample error variance.

## General Conclusions

The statistical tests performed in the preceding section of this chapter have shown that in terms of the sample variance measure the factorial EVOP modification is superior to both the basic Box-Jenkins model and the simplex EVOP modification. However, there is essentially no difference between any of the three forecasting techniques with respect to the average forecast error. From Table 1 it can be seen that in only one of the time series was there a sizeable difference in the average forecast error.

It has been the objective of this investigation to improve a prediction and control procedure, which is already adaptive in nature,

Table 9. Box-Jenkins Model with Simplex EVOP Versus Box-Jenkins Model with Factorial EVOP

[Hypothesis: Simplex EVOP Modification is no worse than Factorial EVOP Modification. $N = 7$; $\alpha = .025$; $T_c = 2$.]

| Time Series | Basic Box-Jenkins Model | Box-Jenkins Model with Factorial EVOP | Difference | Absolute Rank | Positive Rank | Negative Rank |
|---|---|---|---|---|---|---|
| *TEST A.* | *Average Prediction Error* | | | | | |
| 1 | .208 | .147 | .061 | 2 | 2 | - |
| 2 | -2.767 | -.964 | 1.803 | 6 | 6 | - |
| 3 | .440 | 1.997 | -1.557 | 5 | - | 5 |
| 4 | -4.834 | -4.726 | .108 | 3 | 3 | - |
| 5 | -89,046.940 | -.416 | 89,046.524 | 7 | 7 | - |
| 6 | -4.748 | -4.954 | -.206 | 4 | - | 4 |
| 7 | .642 | .643 | -.001 | 1 | - | 1 |
| | | | | | | $T = 10 > T_c = 2$ Accept Hypothesis |
| *TEST B.* | *Sample Error Variance* | | | | | |
| 1 | 58.077 | 44.821 | 13.256 | 4 | 4 | |
| 2 | 114.462 | 84.281 | 30.181 | 5 | 5 | |
| 3 | 236.425 | 223.745 | 12.680 | 3 | 3 | |
| 4 | 211.326 | 215.195 | -3.869 | 2 | - | 2 |
| 5 | $7{,}972{,}055.81 \times 10^4$ | 176.898 | 7,971,878.912 | 7 | 7 | |
| 6 | 8,771.934 | 223.430 | 8,548.504 | 6 | 6 | |
| 7 | 69.431 | 68.622 | .809 | 1 | 1 | |
| | | | | | | $T = 2 = T_c = 2$ Reject Hypothesis |

by making its control parameters self-adaptive to changes in the nature of the time series. It has been shown that the factorial EVOP modification does significantly reduce the error variance while yielding no worse value for the average forecast error. However, to further illustrate from a graphical or visual point of view the performance of the three prediction procedures, an inspection and analysis of Figures 10 through 30 will be made. These graphs show portions of both the actual time series values and the predicted values using the three different procedures.

Time series one is shown on Figures 10 through 12 along with the results of the three prediction procedures. Series one is purely random and consequently all three procedures perform similarly. It can be seen that the factorial EVOP procedure does produce smaller over and under reactions to the noise in the system.

Series two contains a step increase. Figures 13 through 15 show how each of the three prediction systems react to this change in the nature of the time series. It can be noticed that both the simplex and factorial EVOP procedures react to the change and reach the new constant level in fewer time periods than the basic model. However, the factorial EVOP procedure reaches the new constant level in fewer time periods than even the simplex EVOP procedure.

Figures 16 through 18 show the response of the three forecasting procedures to an impulse in the time series. Because of the adaptive nature of the three techniques, a one period lag is exhibited by all three. However, the basic Box-Jenkins model and the factorial EVOP procedure result in several periods of large error fluctuations, while the simplex procedure settles back down rather quickly. The reason for this may be explained simply. In the simplex procedure the control parameters adapt themselves each period to changes in the time series. Thus the impulse signals a parameter change one period and the return to the previous level signals another parameter change. However, in the factorial EVOP procedure, the impulse signals an upward shift in a control parameter but several periods are then required to re-evaluate the range estimate of the standard deviation so that a downward parameter shift can occur. Thus during these periods the factorial EVOP procedure and the basic model behave in essentially the same way, depending only upon the adaptive nature of the basic model instead of parameter modification.

Series four, which contains a linear trend component, is presented along with the predicted series in Figures 19 through 21. It can be seen that in this series there is very little difference in the performance of the three prediction procedures. The adaptive nature of the basic Box-Jenkins model more or less makes parameter modification unnecessary because of the manner in which it follows the linear trend in the series.

Time series five, which contains a basic sinusoid or periodic pattern, clearly demonstrated the superiority of the factorial EVOP procedure. In Figures 22 through 24 it can be seen that the basic model and the simplex EVOP modification result in very large prediction errors. In fact, these errors become increasingly larger throughout the remainder of the series values. Complete plots of these two predicted series were not made for reasons of graphical scaling problems. However, observation of Figure 23 reveals that the factorial EVOP procedure with its statistical control limits allows parameter modifications which result in a predicted series which follows closely the periodic peaks and troughs of the series.

Series six contains both a trend component and a periodic or sinusoid component. Although it is not shown on Figures 25 and 27, the errors resulting from the basic Box-Jenkins model and the simplex EVOP modification grow progressively larger throughout the series and lead to the large error variances shown in Table 3. On the other hand, the factorial EVOP procedure again allows parameter modifications which result in a relatively accurate predicted series. It would seem that the continuous parameter changes inherent in the simplex EVOP procedure introduce additional error in series exhibiting periodic behavior.

Series seven, which is very highly autocorrelated, is shown in Figures 28 through 30. Both from a visual inspection and from the statistical calculations, it can be seen that there is very little difference in the performance of the three prediction procedures for this time series.

In conclusion it can be stated that the factorial EVOP modification to the Box-Jenkins prediction and control model yields a more accurate prediction than the basic Box-Jenkins model. Thus the objectives of this investigation have been attained in that:

1. It has been shown that the application of Evolutionary Operation techniques can improve a well-known prediction and control model.

2. It has been shown that the factorial EVOP procedure is superior to the simplex EVOP procedure for the model and time series tested.

Therefore, it can be concluded that the application of the factorial EVOP procedure to the Box-Jenkins model in industrial processes may result in more accurate prediction and control.

## Recommendations for Further Research

There are three areas of further research that can be effectively brought forth from this thesis. The first two of these areas involve the extension of the Box-Jenkins model itself, while the third involves a further investigation of the effects of evolutionary operation procedures on predictive and control models.

The first area for further research involves extending the predictive range of the Box-Jenkins model to more than one time period into the future. As the model now exists, it is capable of predicting a time series value for only a single future period. In many prediction and control applications, it is of value to be able to predict for several future periods. Most of the other prediction models mentioned in Chapter I of this thesis do have a multi-period predictive capability. Therefore, if this capability could also be developed for the Box-Jenkins model, then its performance could be compared with the performance of other models to determine if an improved prediction system might result.

The second area for further research involves expanding the practical prediction model described in Chapter II to more than three terms. The additional terms are already defined in the general form of the prediction and control model. As yet, however, little research has been performed to determine if the inclusion of some of these additional terms will improve the prediction and control model.

The third and final area of further research involves an investigation to determine the effect of the evolutionary operation procedures on time series with varying magnitudes of the random components. In this thesis the random components had a relatively small magnitude with respect to the basic nature of the time series used. Under these conditions it was shown that the factorial EVOP procedure could yield an improved performance of the Box-Jenkins model over that obtained by the use of the simplex EVOP procedure. Limited experimentation with time series possessing a small signal to random noise ratio has indicated that the simplex procedure tends to yield a better prediction. An investigation to determine the relationships between the EVOP procedures and the magnitude of the random noise using both the Box-Jenkins model and other prediction model would yield valuable information.

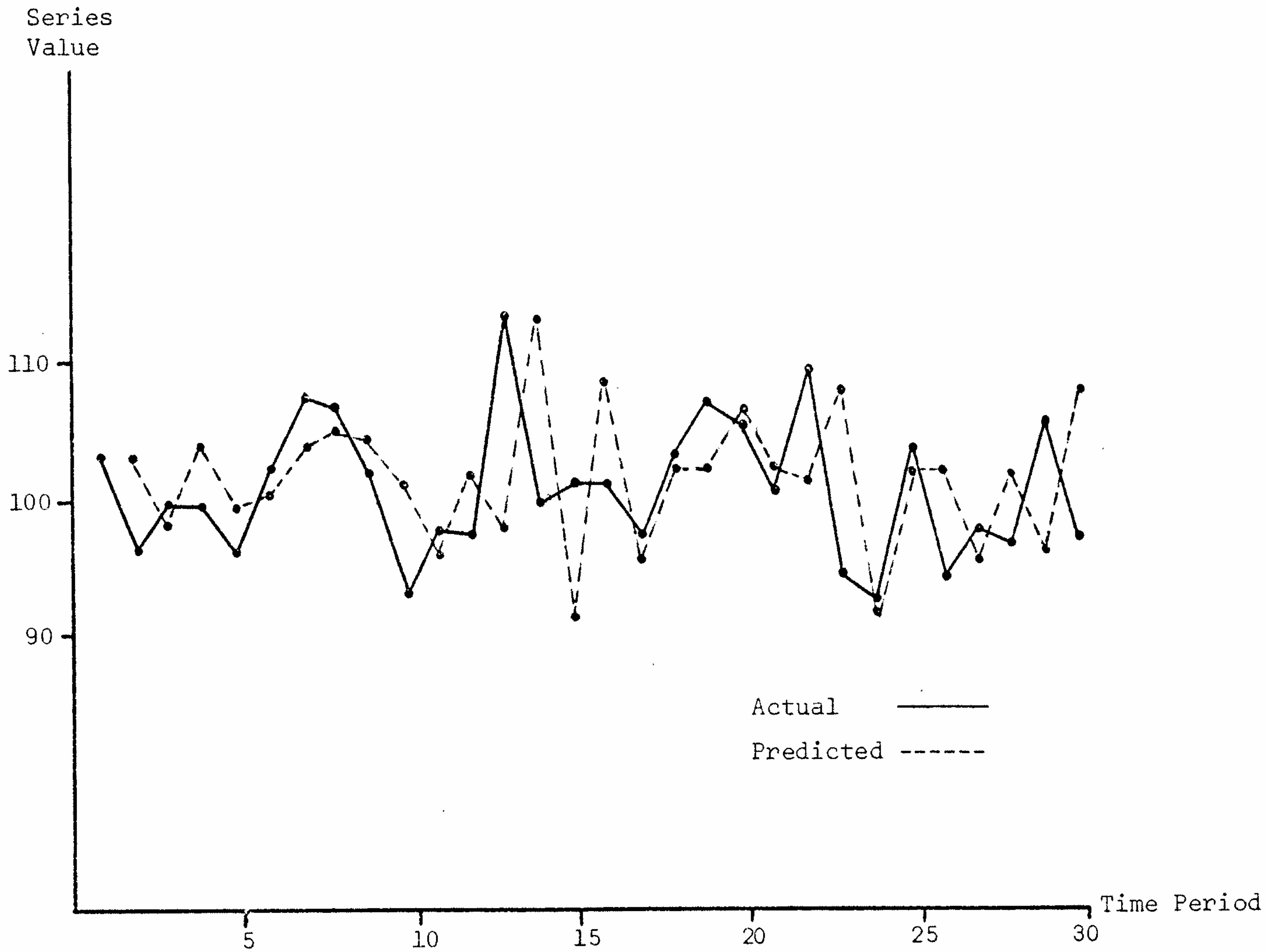


Figure 10. Performance of Basic Box-Jenkins Model on Series One

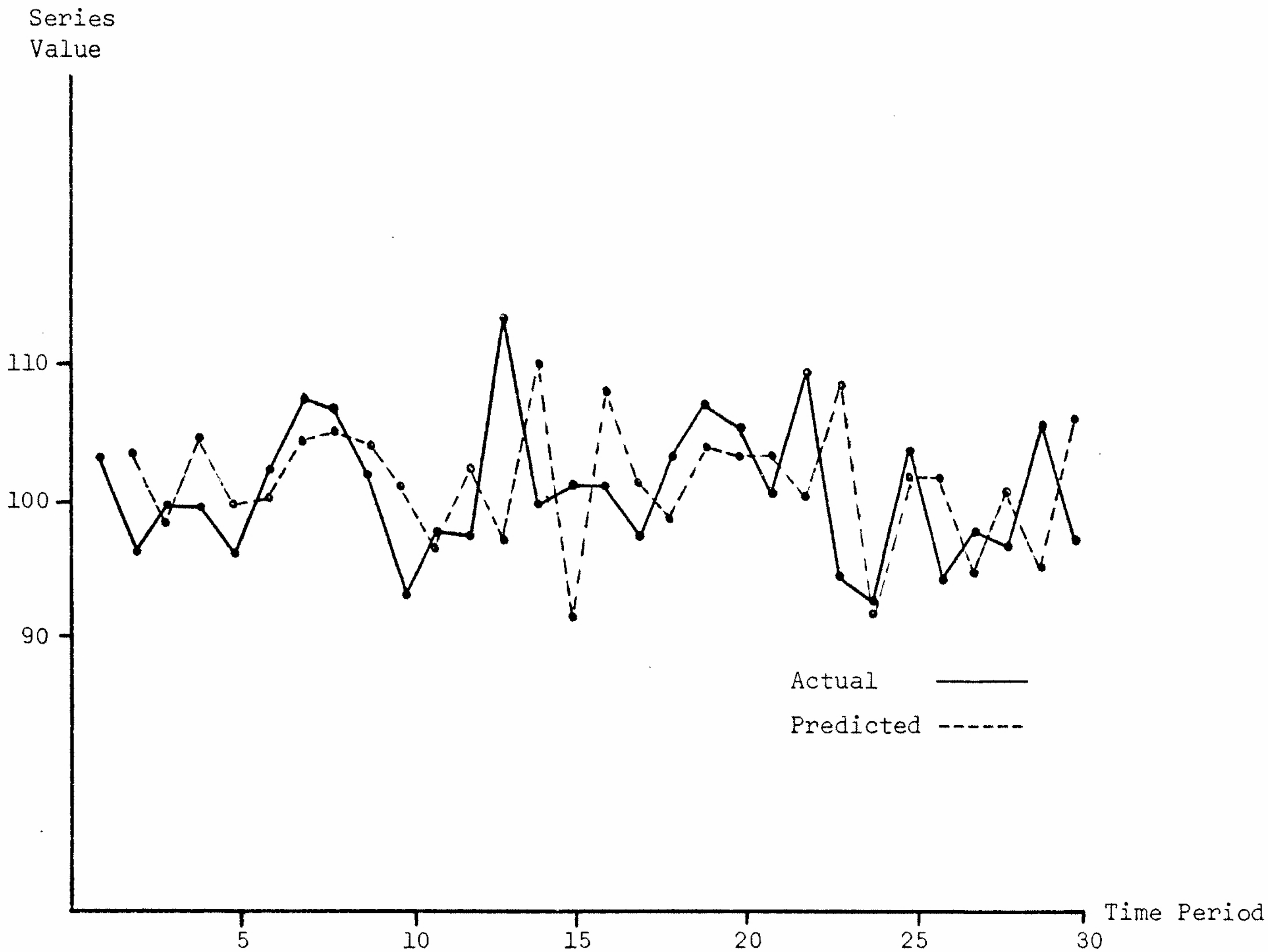


Figure 11. Performance of Box-Jenkins Model with Factorial EVOP on Series One

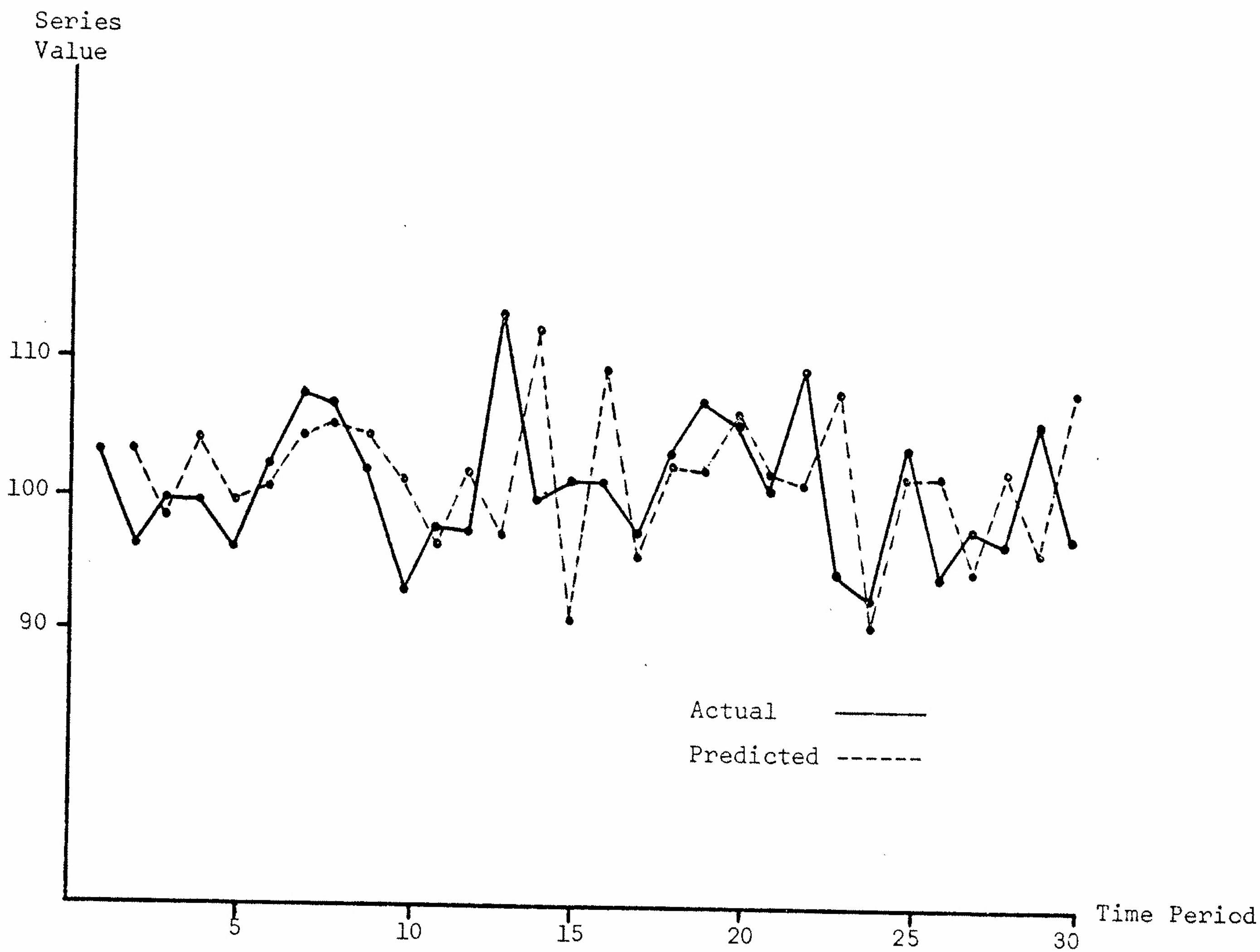


Figure 12. Performance of Box-Jenkins Model with Simplex EVOP on Series One

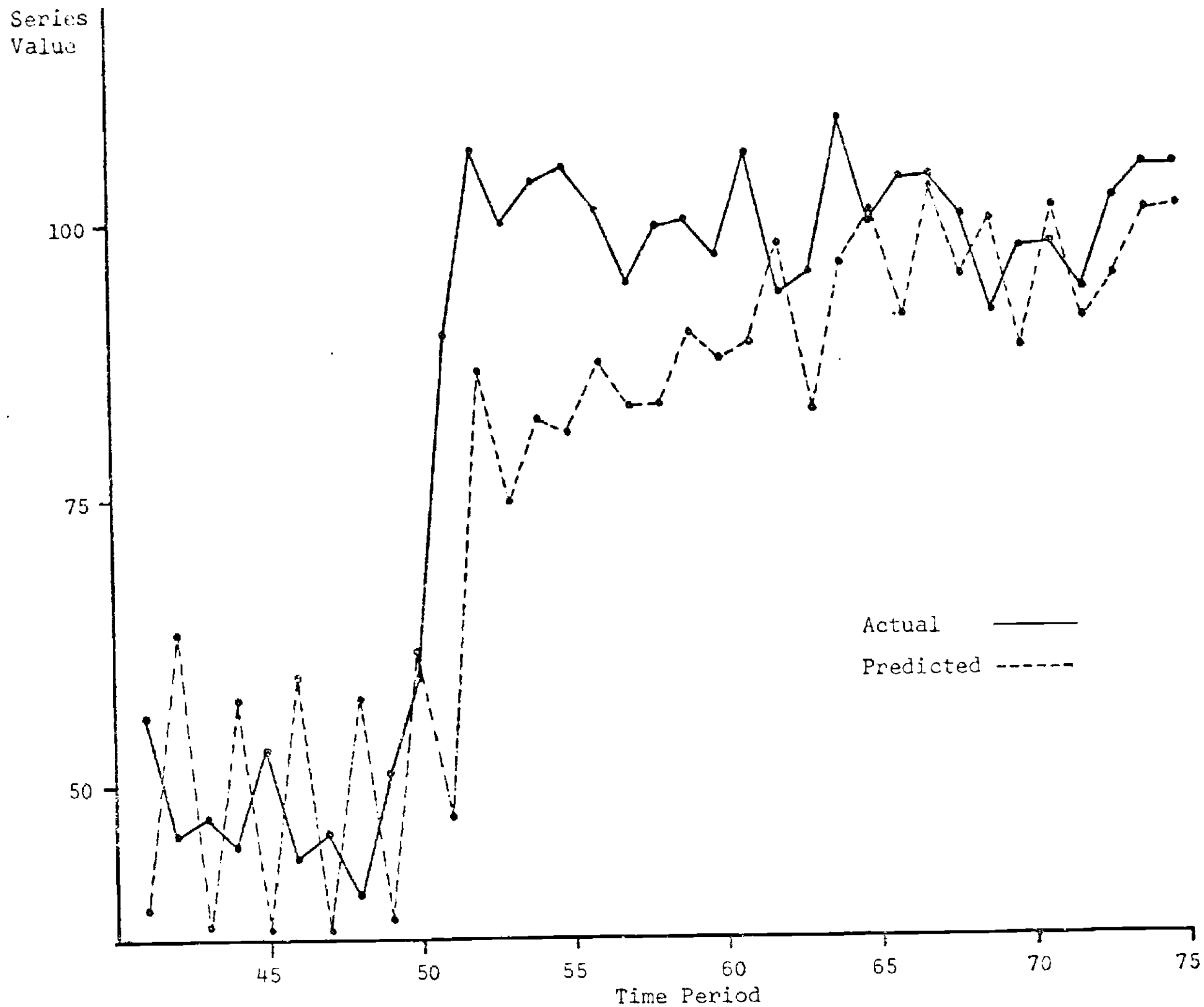


Figure 13. Performance of Basic Box-Jenkins Model on Series Two

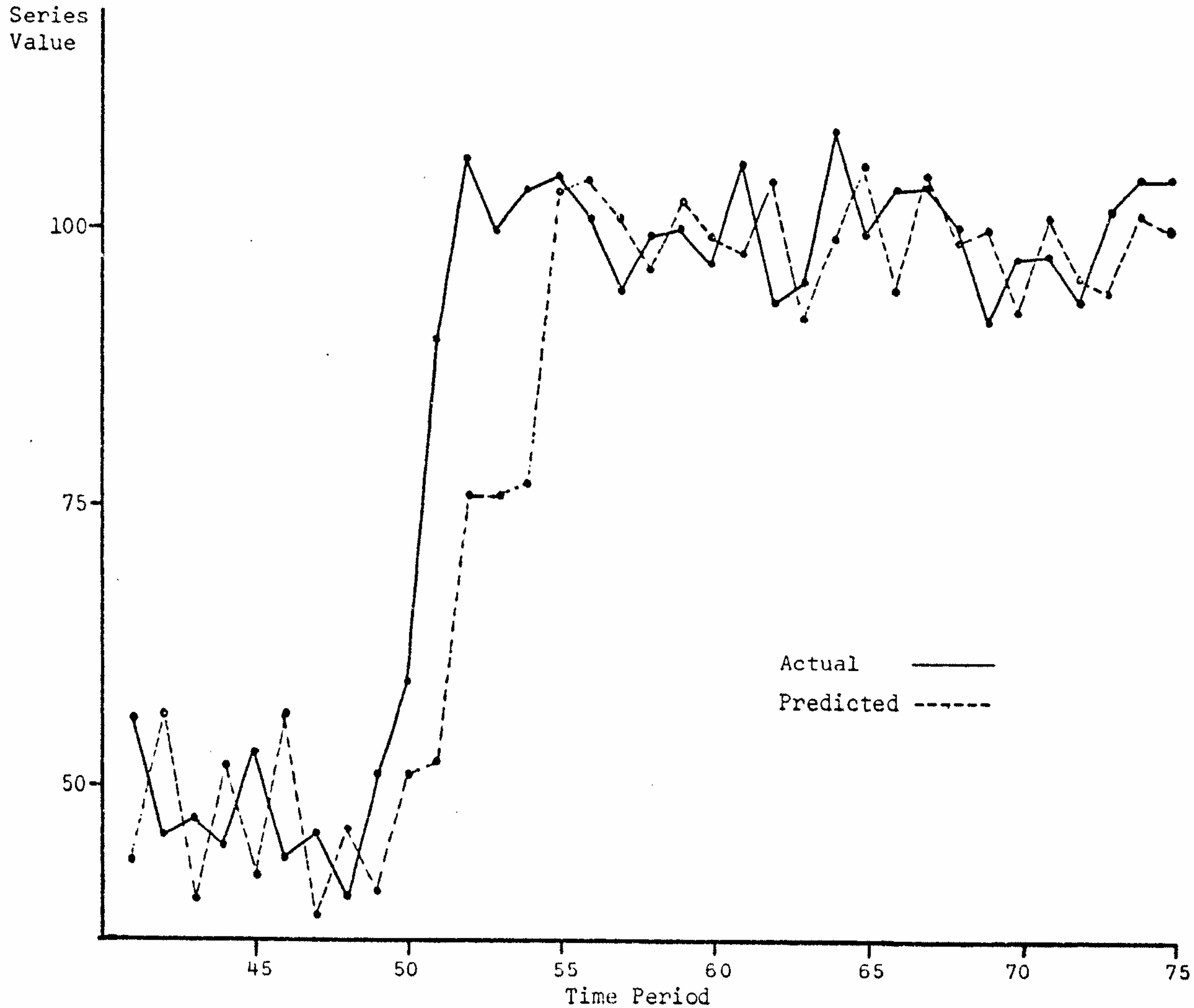


Figure 14. Performance of Box-Jenkins Model with Factorial EVOP on Series Two

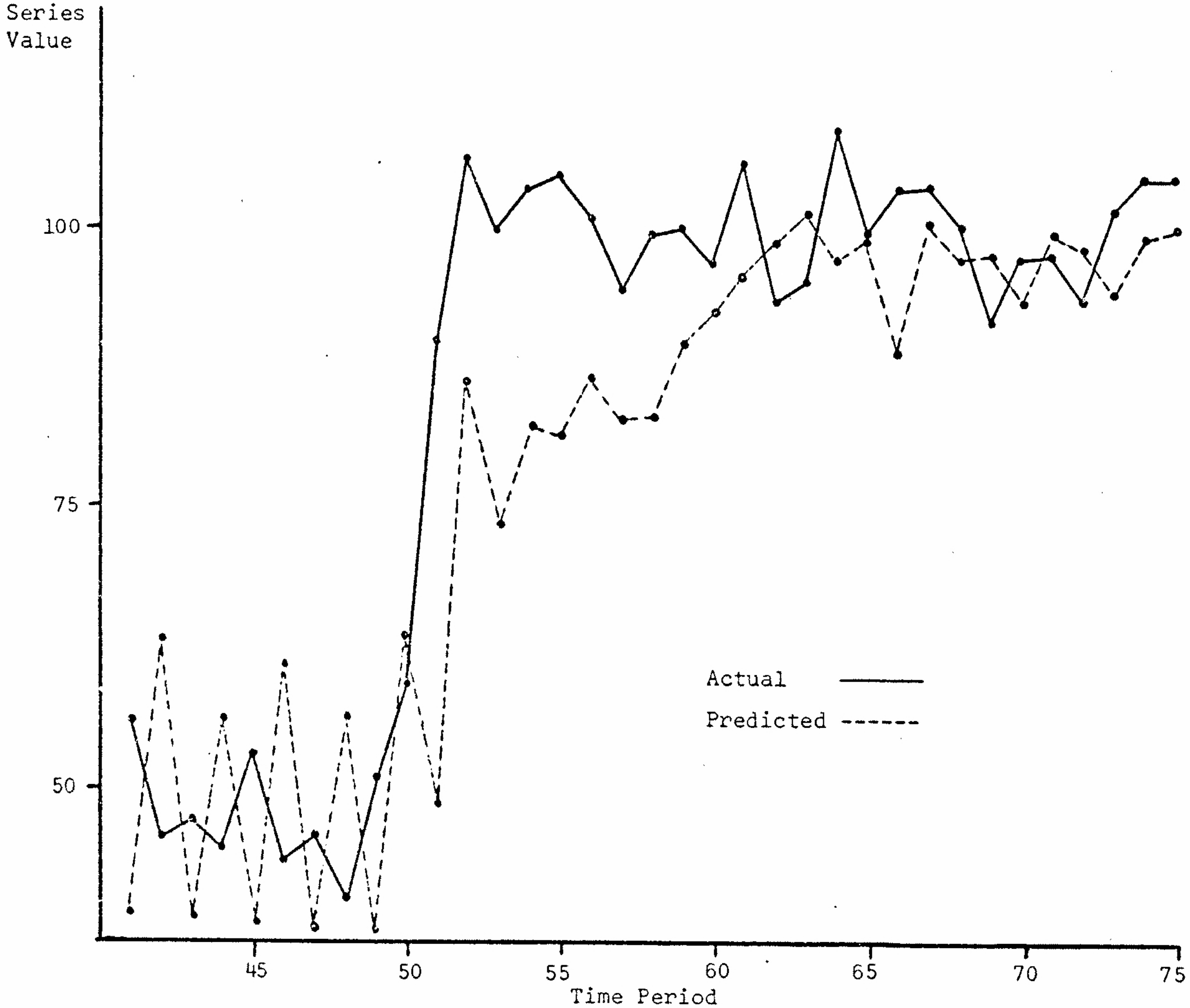


Figure 15. Performance of Box-Jenkins with Simplex EVOP on Series Two

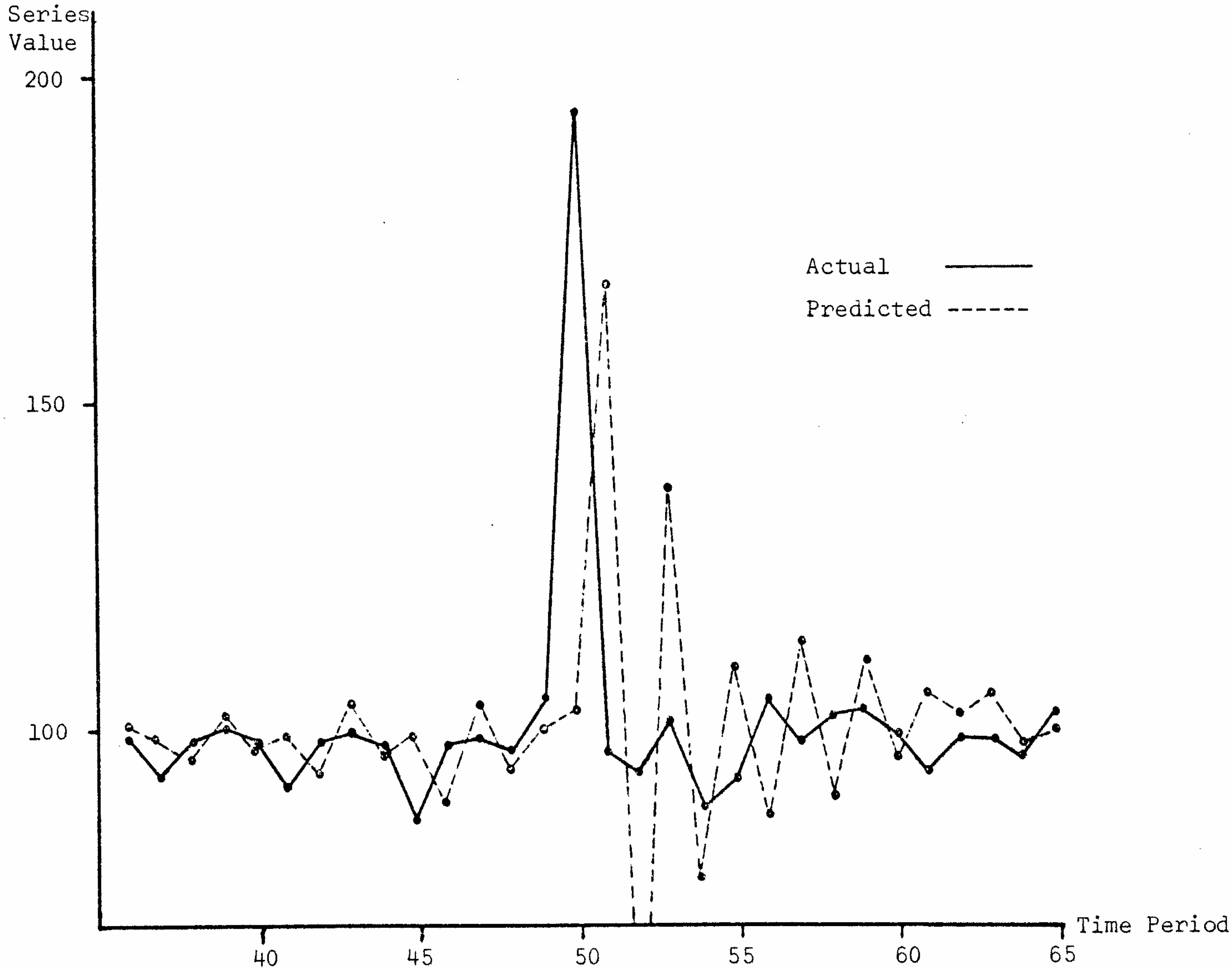


Figure 16. Performance of Basic Box-Jenkins Model on Series Three

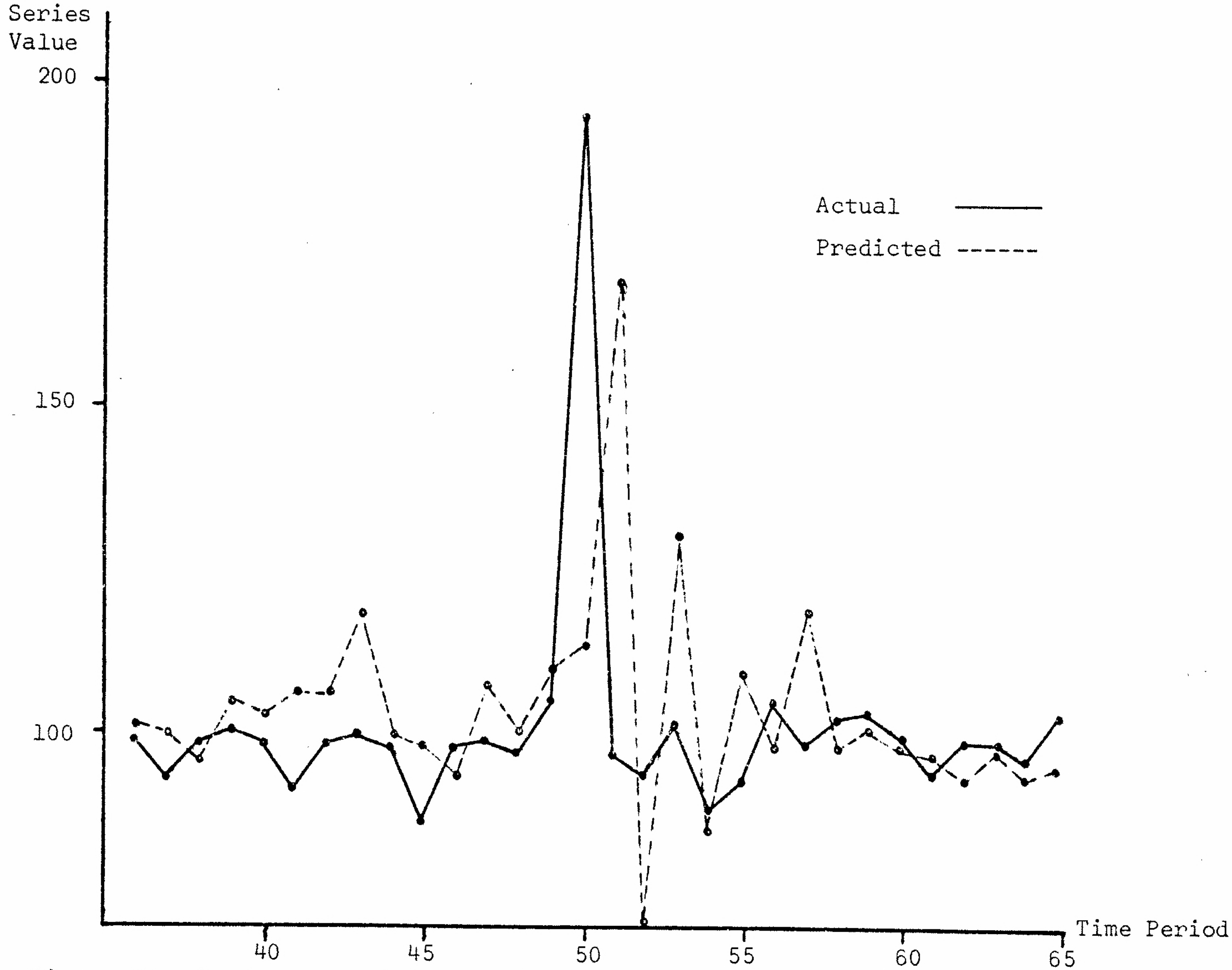


Figure 17. Performance of Box-Jenkins Model with Factorial EVOP on Series Three

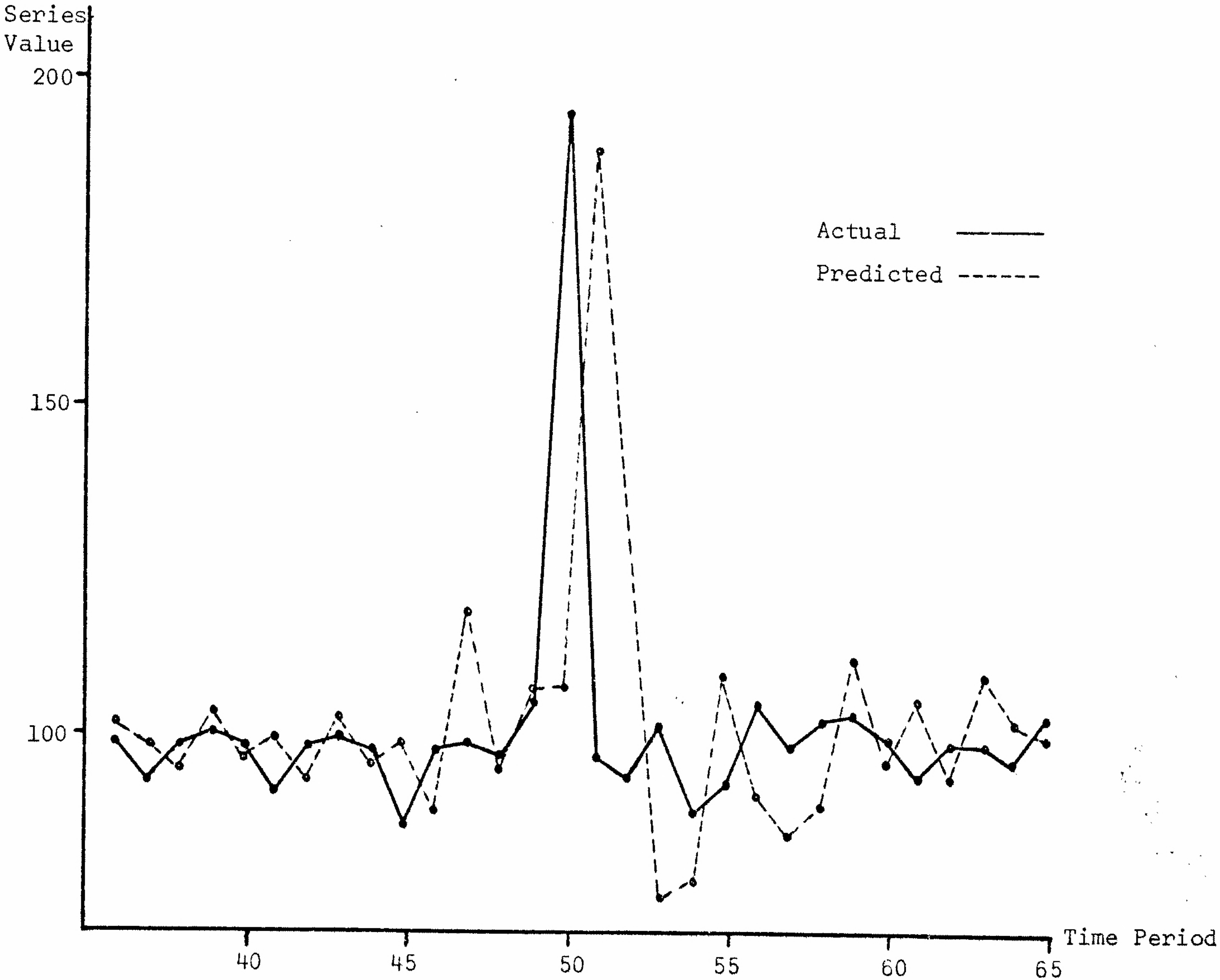

Figure 18. Performance of Box-Jenkins Model with Simplex EVOP on Series Three

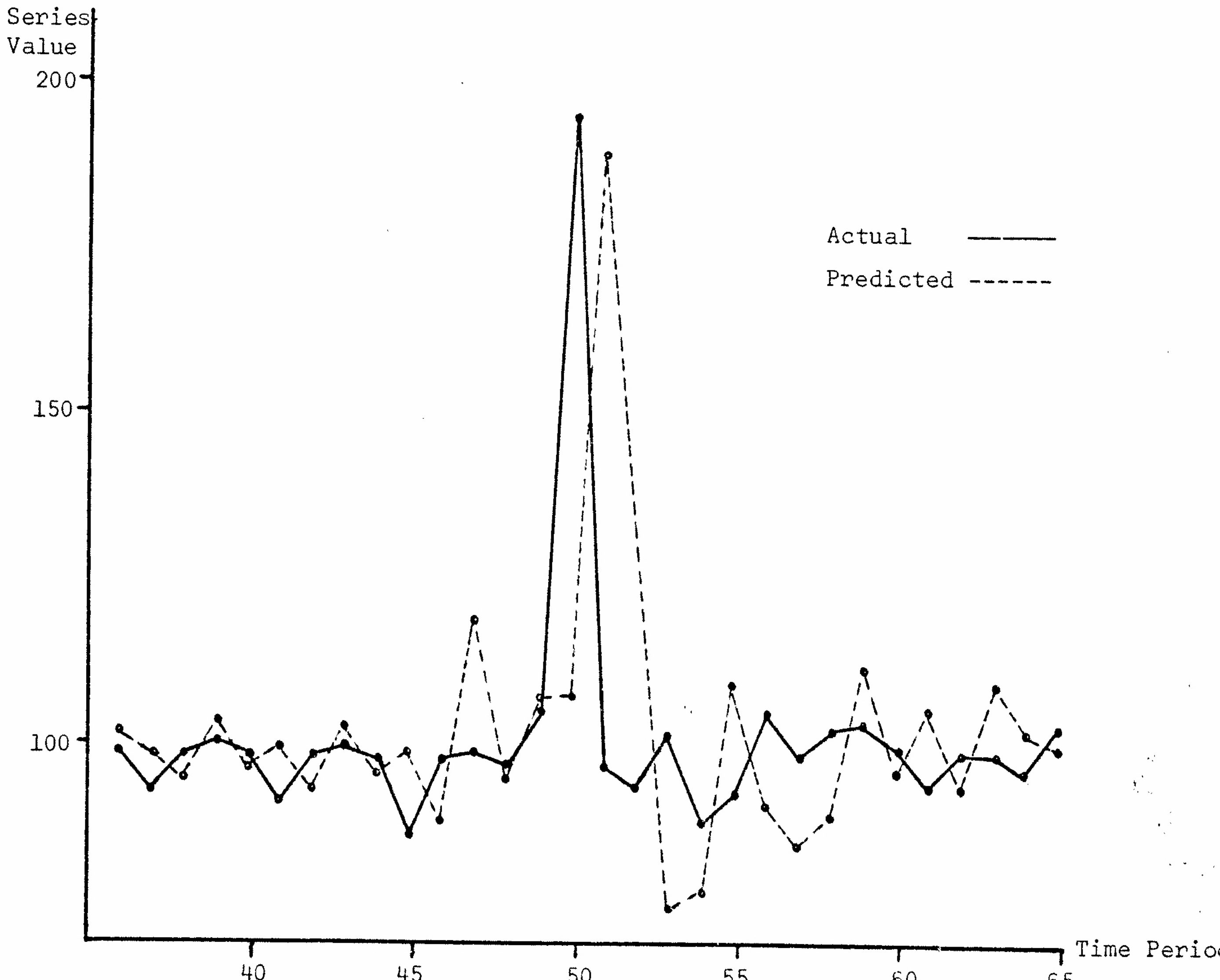

Figure 18. Performance of Box-Jenkins Model with Simplex EVOP on Series Three

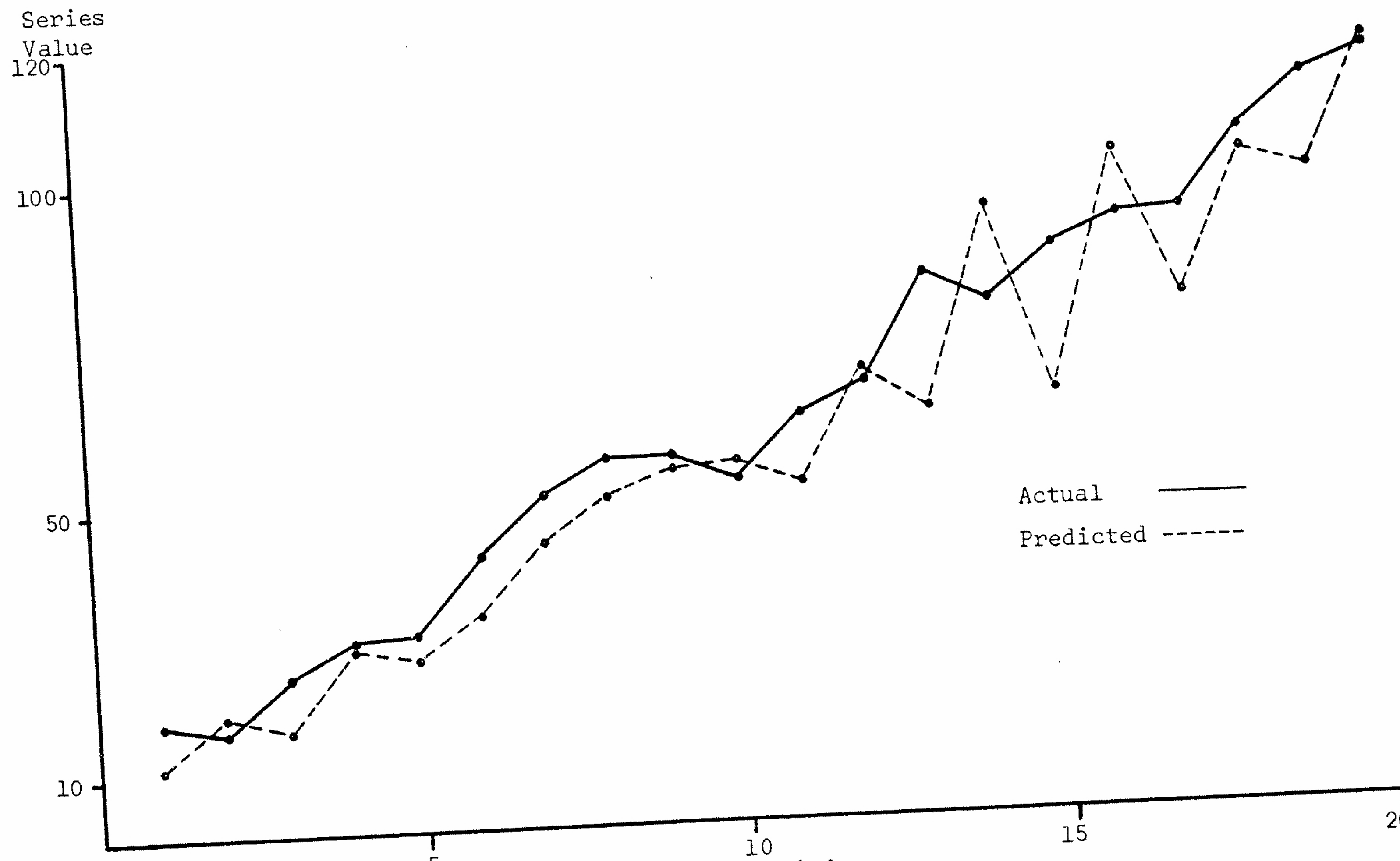


Figure 19. Performance of Basic Box-Jenkins Model on Series Four

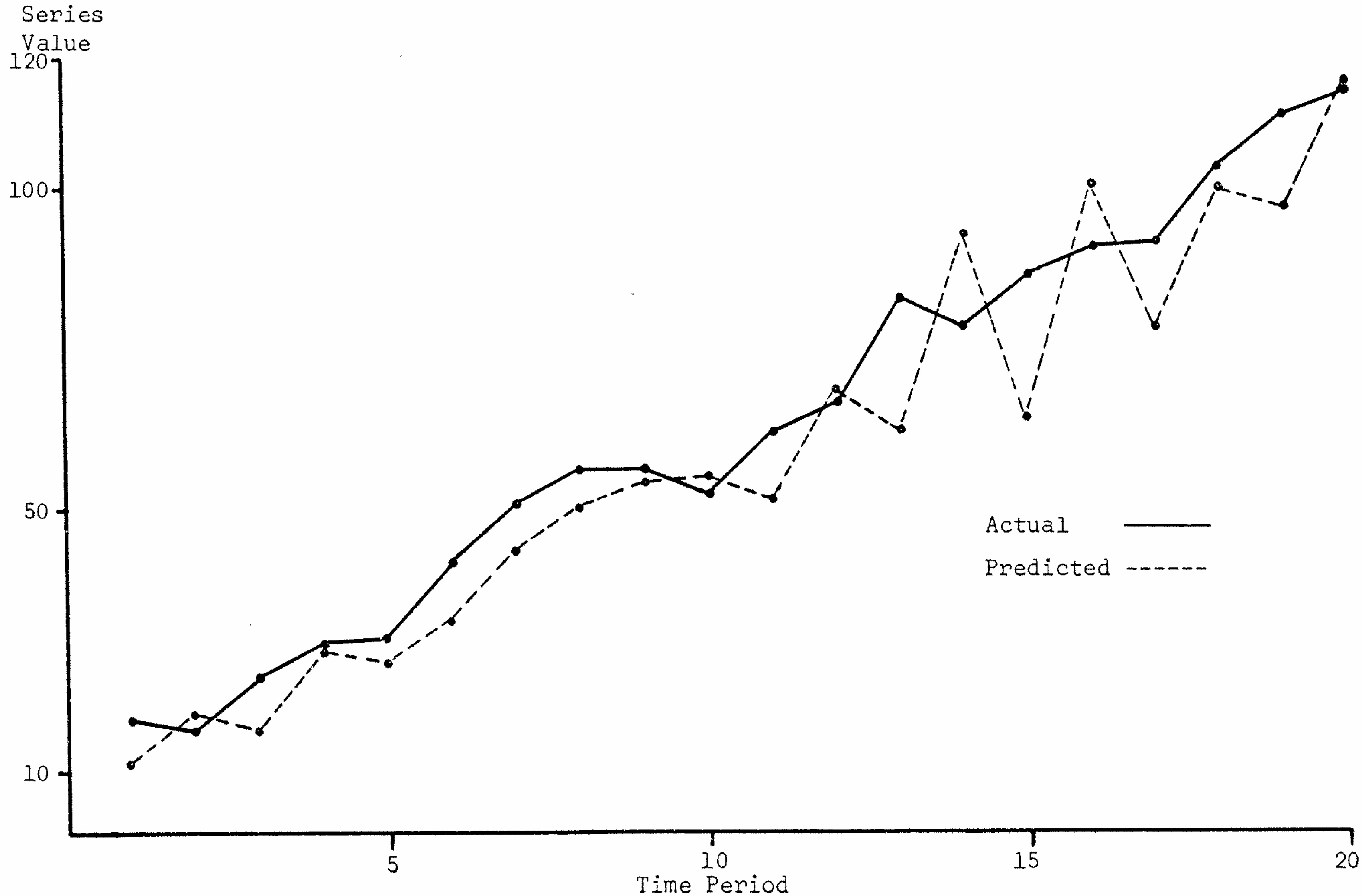

Figure 19.  Performance of Basic Box-Jenkins Model on Series Four

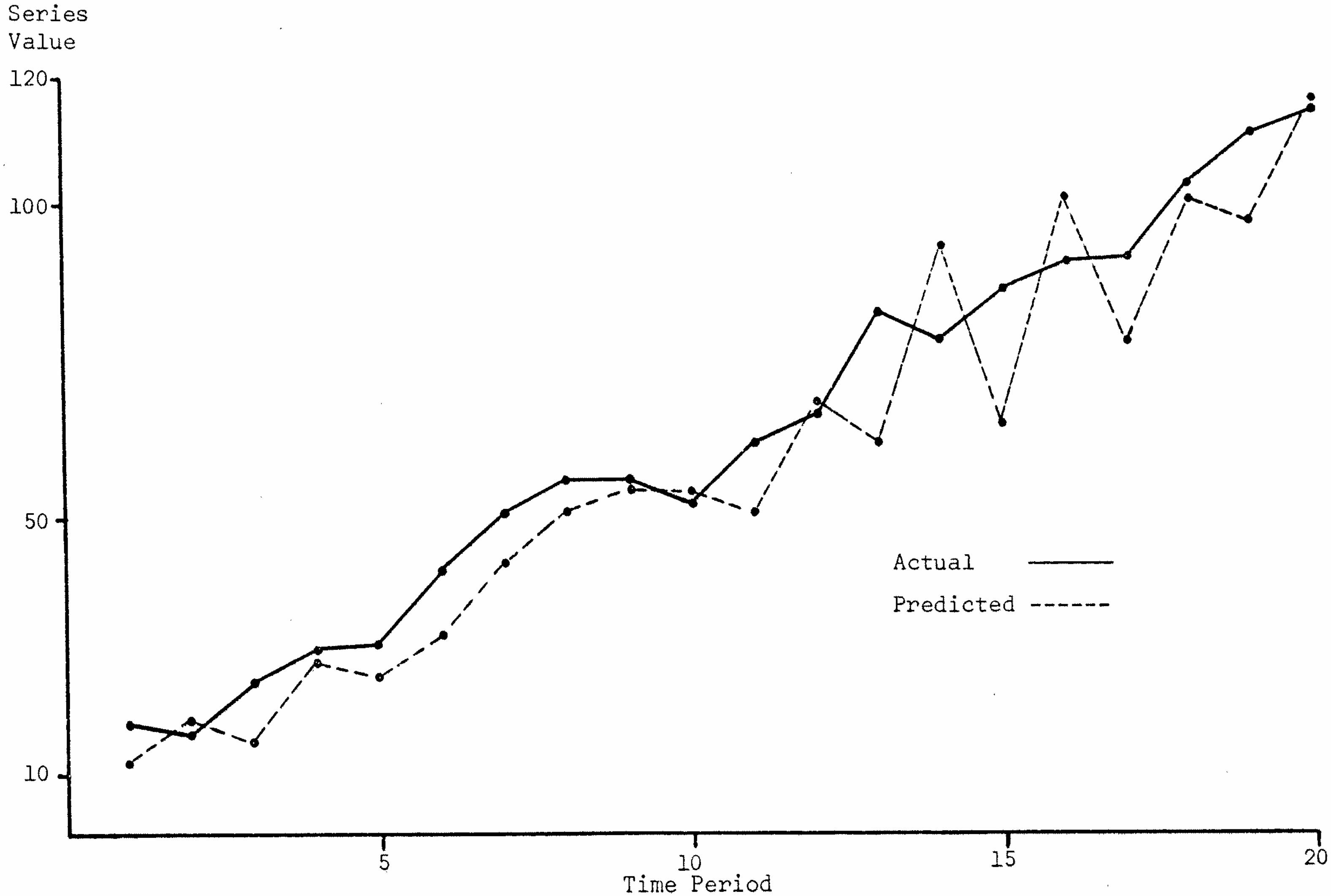


Figure 20. Performance of Box-Jenkins Model with Factorial EVOP on Series Four

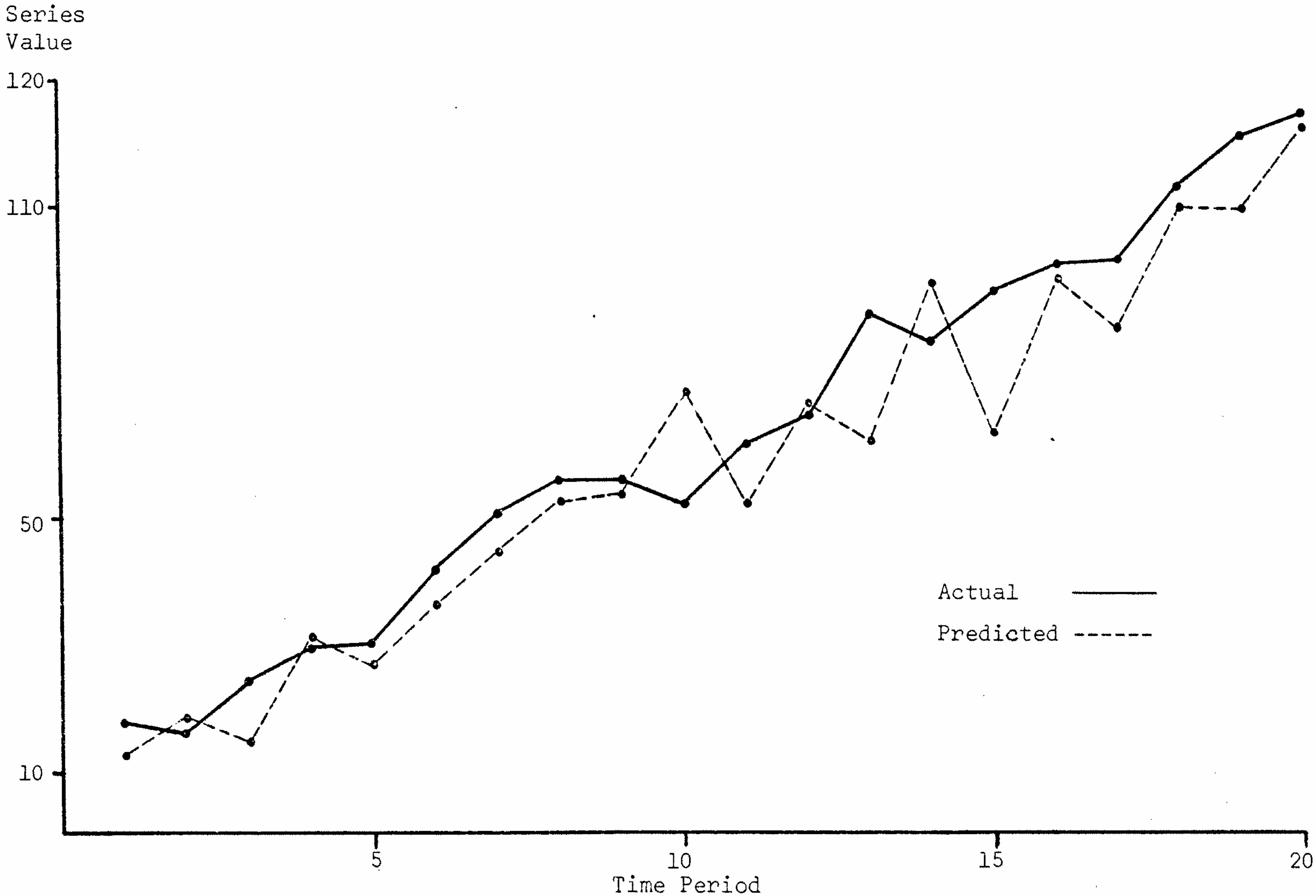


Figure 21. Performance of Box-Jenkins Model with Simplex EVOP on Series Four

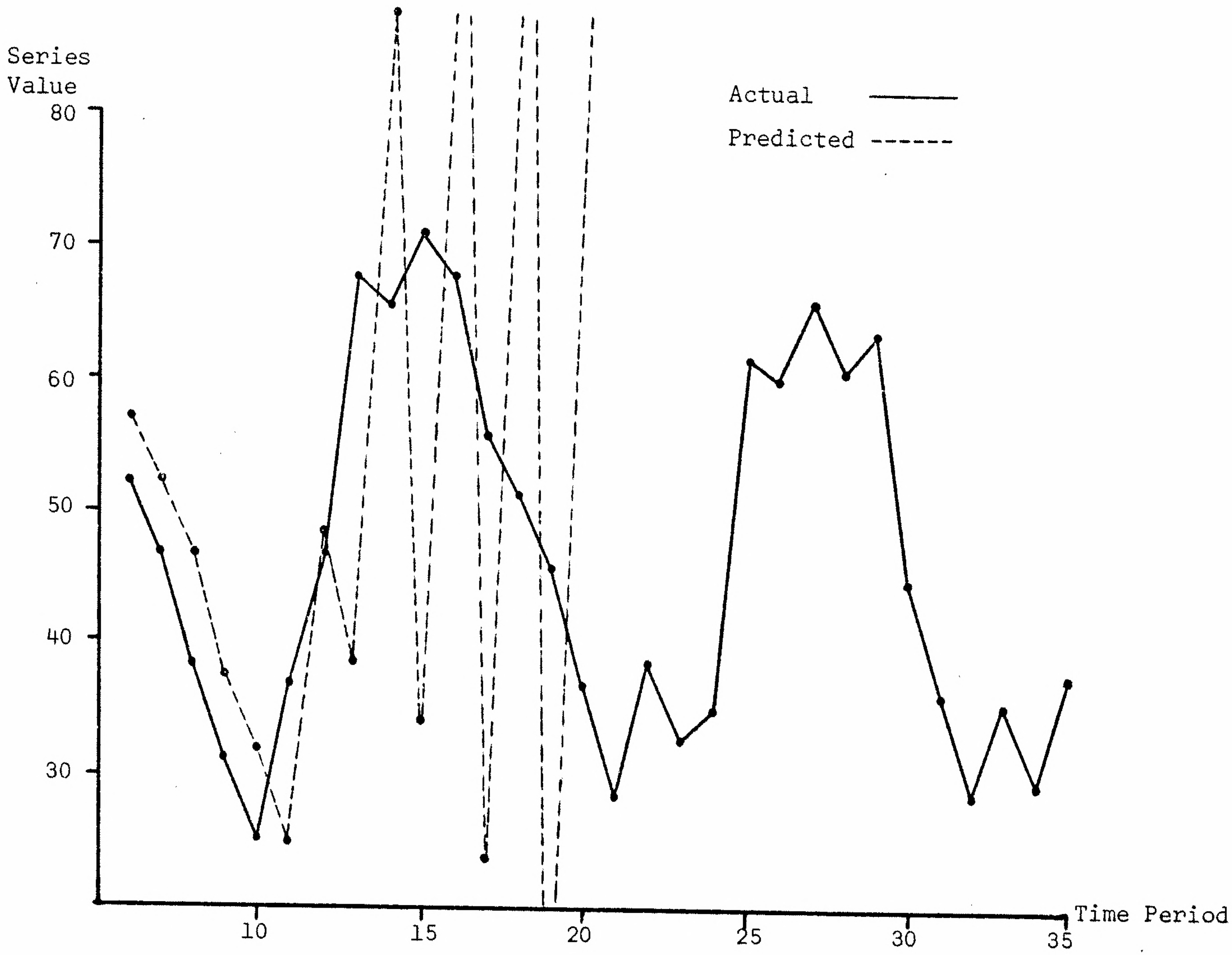


Figure 22. Performance of Basic Box-Jenkins Model on Series Five

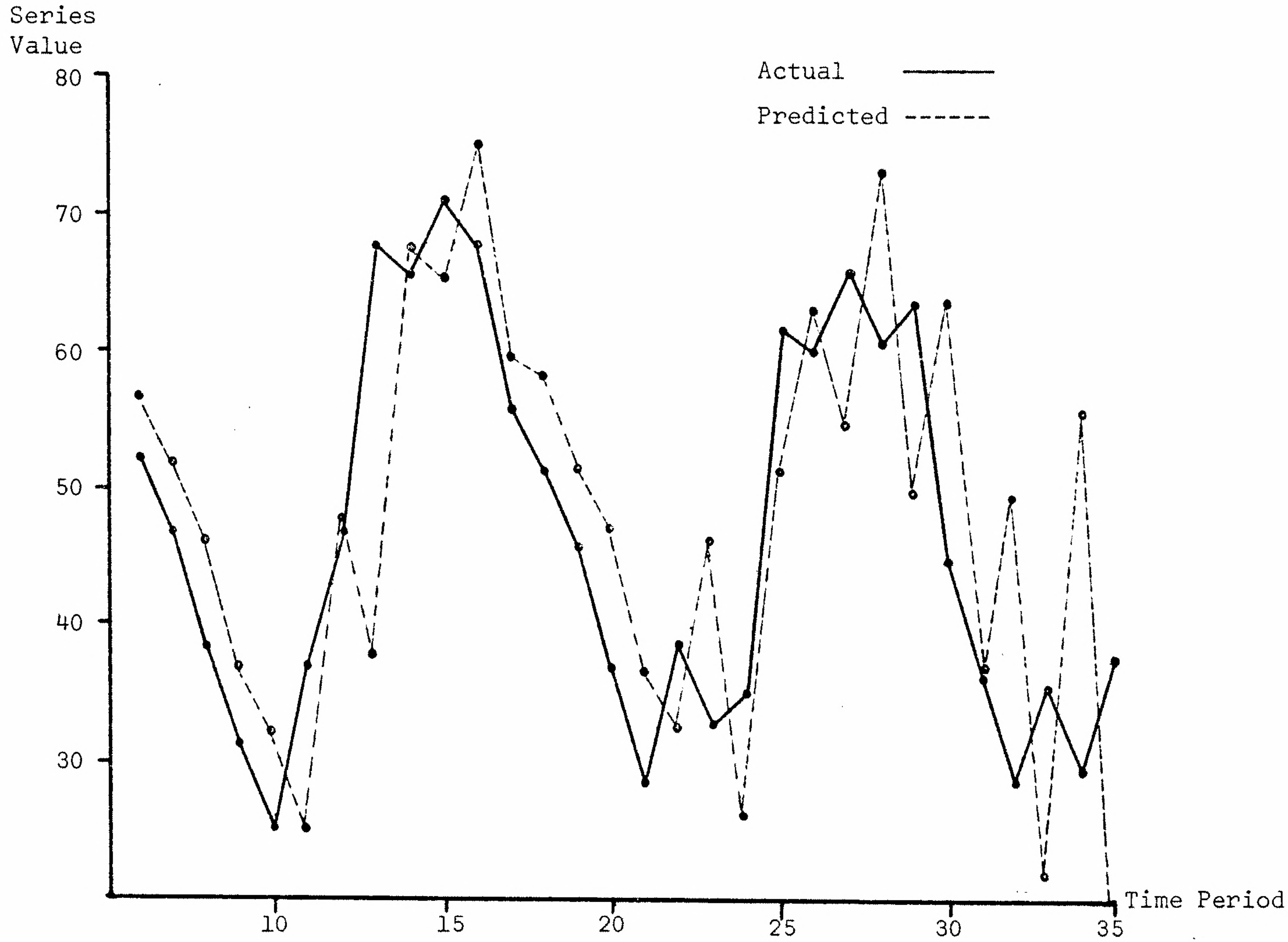


Figure 23. Performance of Box-Jenkins Model with Factorial EVOP on Series Five

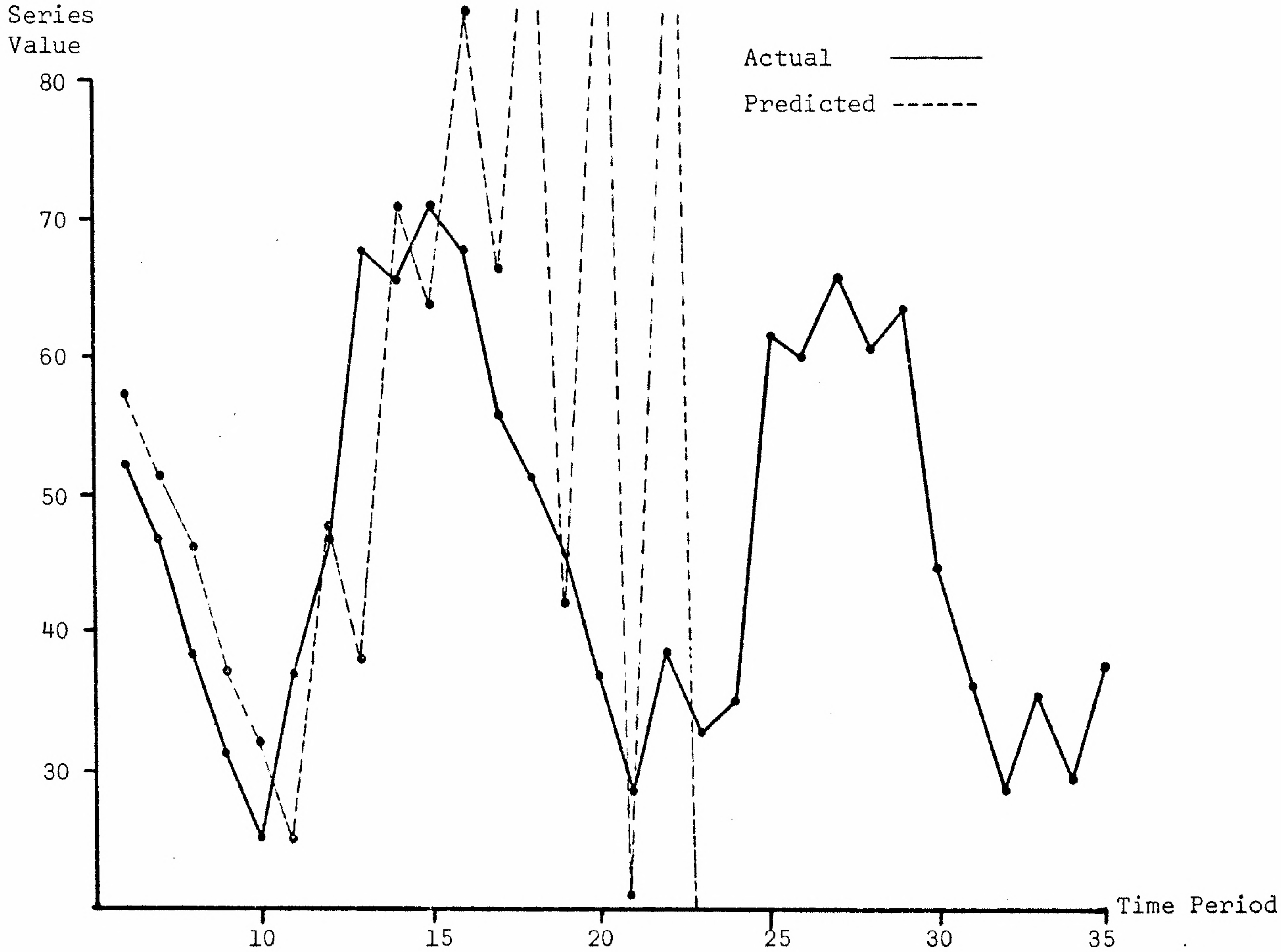


Figure 24. Performance of Box-Jenkins Model with Simplex EVOP on Series Five

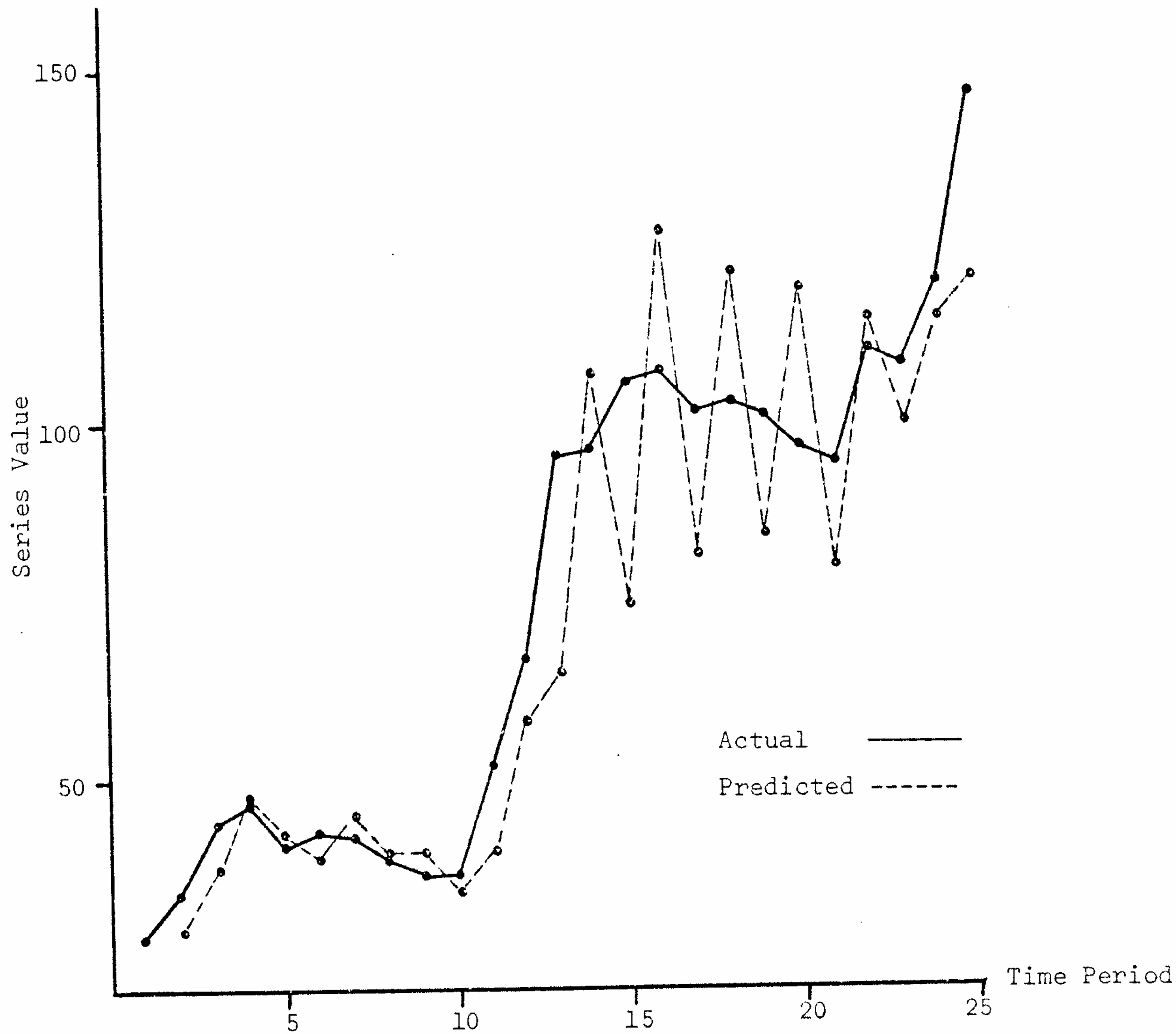


Figure 25. Performance of Basic Box-Jenkins Model on Series Six

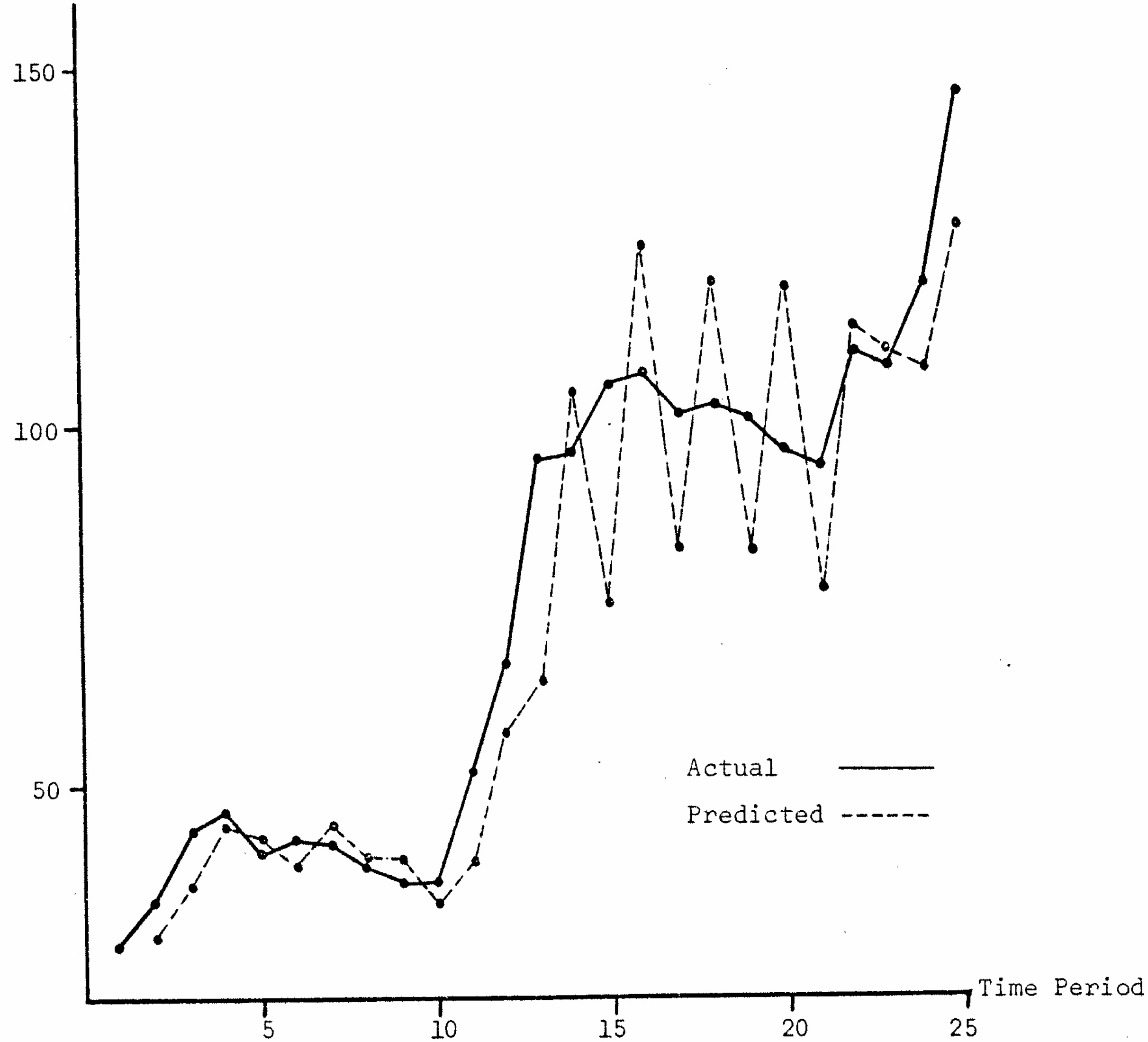


Figure 26. Performance of Box-Jenkins Model with Factorial EVOP on Series Six

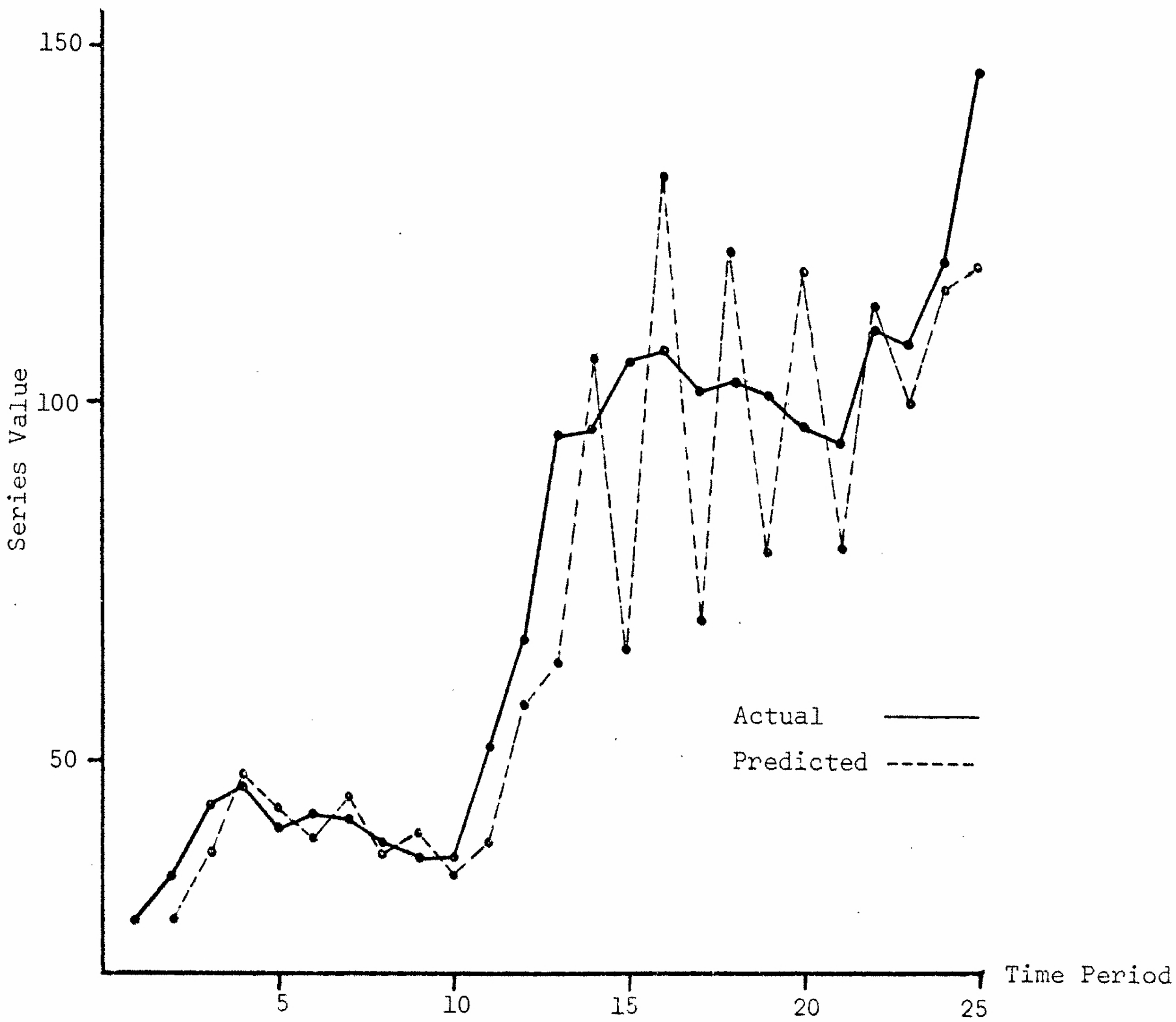


Figure 27. Performance of Box-Jenkins Model with Simplex EVOP on Series Six

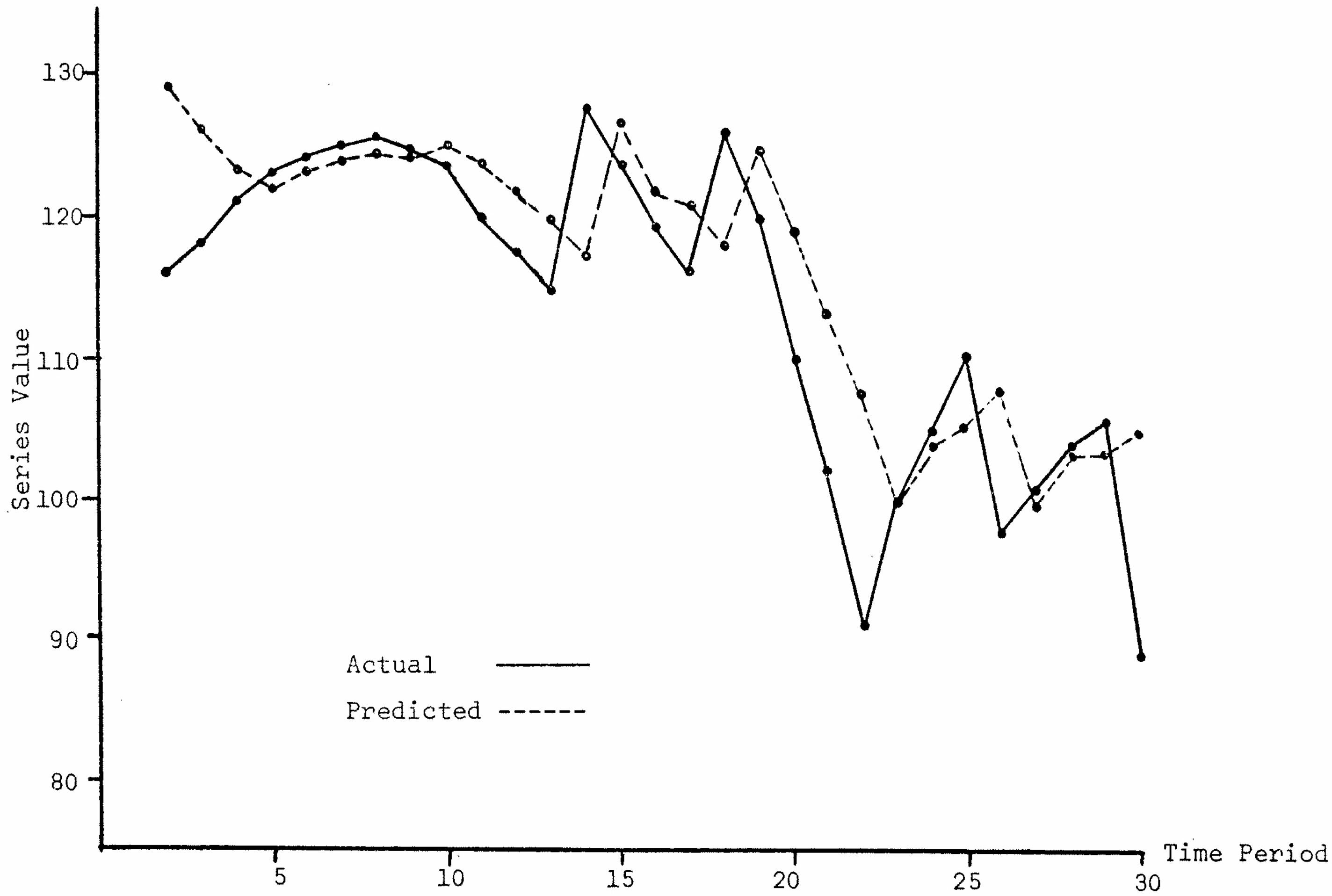


Figure 28. Performance of Basic Box-Jenkins Model on Series Seven

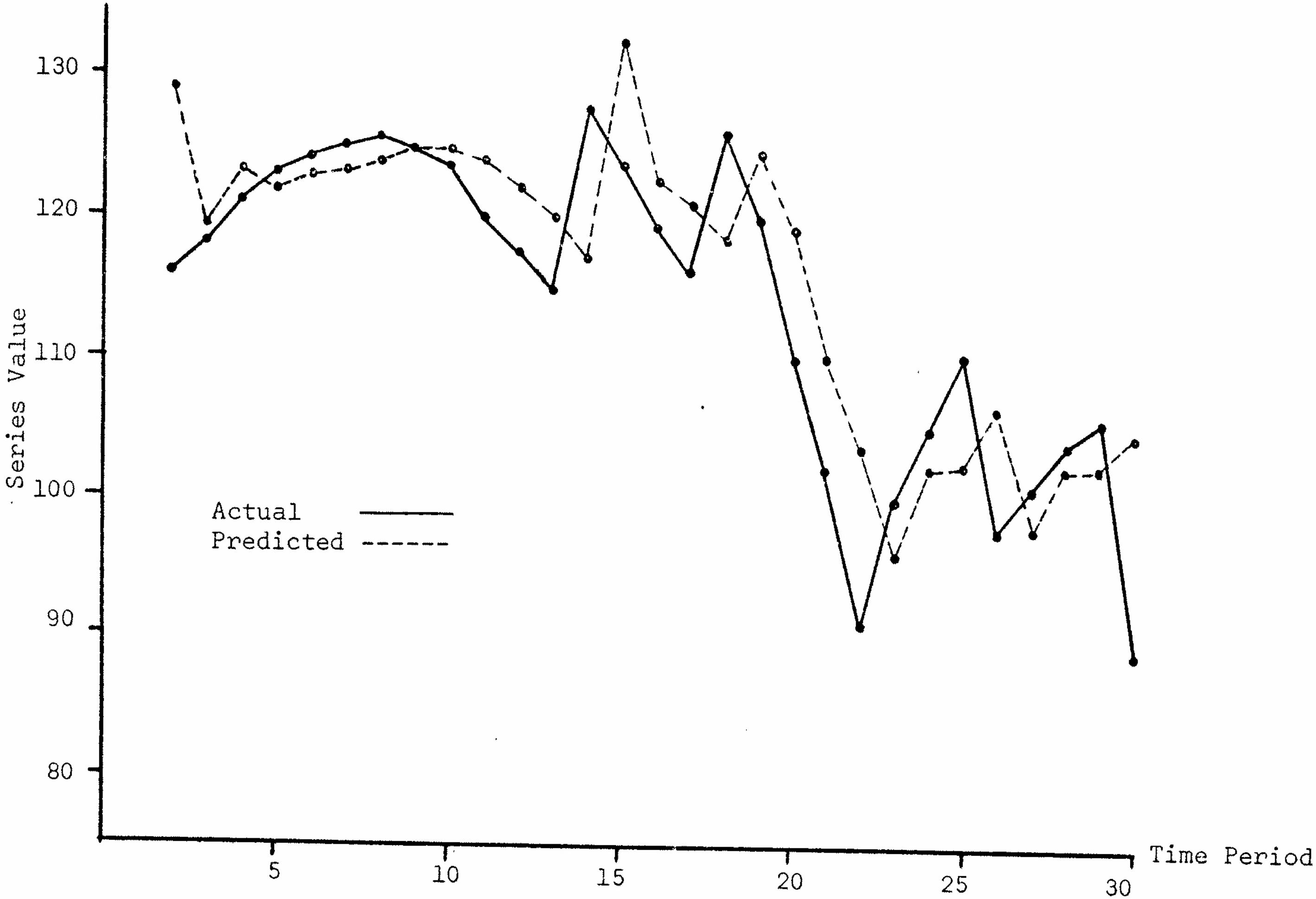


Figure 29. Performance of Box-Jenkins Model with Factorial EVOP on Series Seven

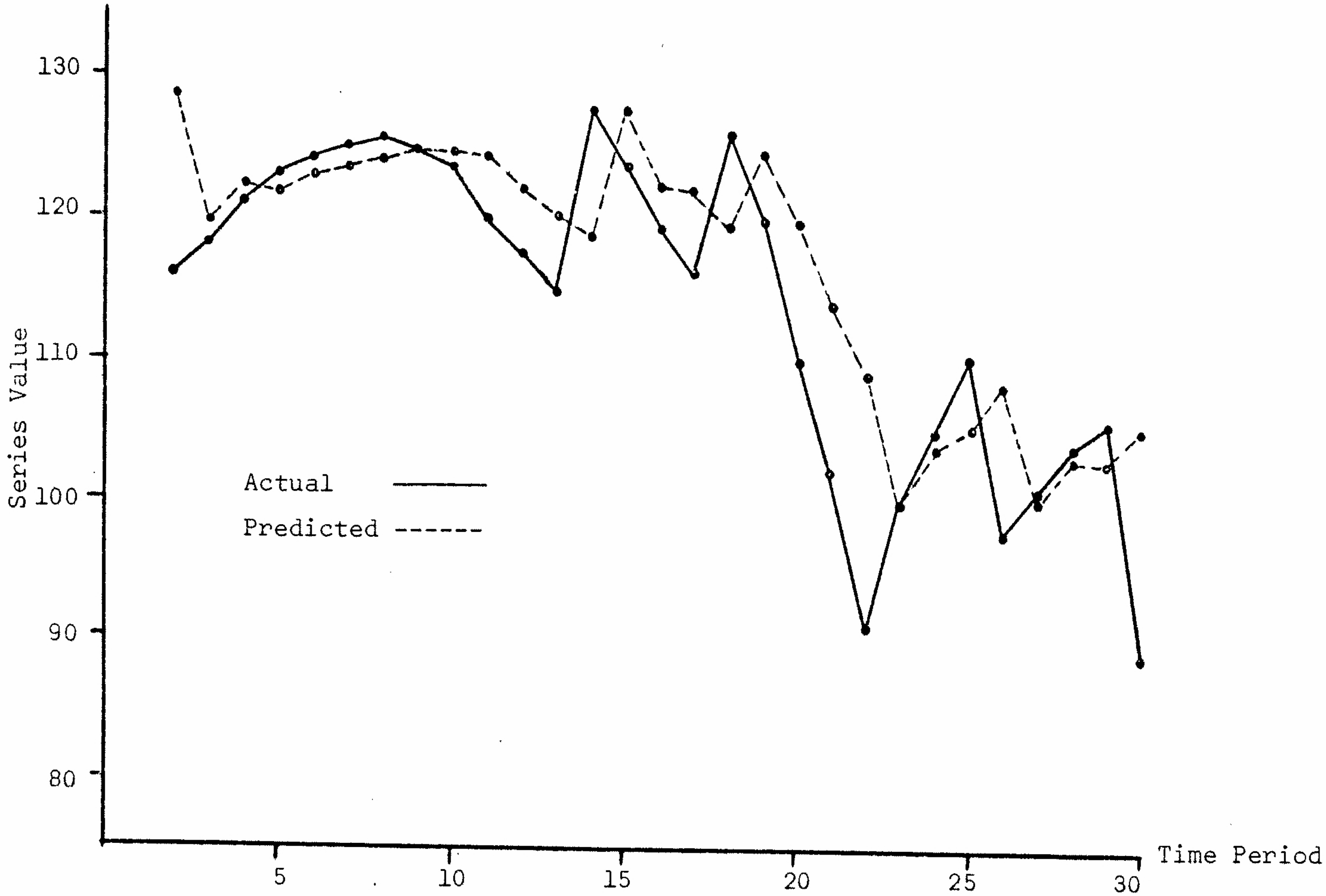


Figure 30. Performance of Box-Jenkins Model with Simplex EVOP on Series Seven

# BIBLIOGRAPHY


1. Box, G. E. P., "Evolutionary Operation: A Method for Increasing Industrial Productivity," *Applied Statistics*, Vol. 6, No. 2 (June, 1957), pp. 81-101.

2. Box, G. E. P., and Draper, N. R., *Evolutionary Operation*, John Wiley & Sons, Inc., New York, 1969.

3. Box, G. E. P., and Hunter, J. S., "Condensed Calculations for Evolutionary Operation," *Technometrics*, Vol. 1, No. 1 (Feb., 1959), pp. 77-95.

4. Box, G. E. P., and Jenkins, G. M., "Some Statistical Aspects of Adaptive Optimization and Control," *Journal of the Royal Statistical Society*, B24, No. 2 (1962), pp. 297-332.

5. Brown, R. G., *Decision Rules for Inventory Management*, Holt, Rinehart & Winston, Inc., New York, 1967.

6. Brownlee, K. A., *Statistical Theory and Methodology in Science and Engineering*, 2nd Edition, John Wiley & Sons, Inc., New York, 1965.

7. Burgess, John T., "Adaptive Forecasting," Presented at the 13th Annual TIMS Meeting, Philadelphia, Penn., Sept. 6-8, 1966.

8. Burr, I. W., *Engineering Statistics and Quality Control*, McGraw-Hill, New York, 1953.

9. Chow, W. M., "Adaptive Control of the Exponential Smoothing Constant," *Journal of Industrial Engineering*, Vol. 16, No. 5 (Sept.-Oct. 1965).

10. Hicks, C. R., *Fundamental Concepts in the Design of Experiments*, Holt, Rinehart and Winston, Inc., New York, 1964.

11. Montgomery, D. C., "Adaptive Control of Exponential Smoothing Parameters by Evolutionary Operation," *AIIE Transactions*, Vol. 11, No. 3 (September, 1970).

12. Ostle, B., *Statistics in Research*, The Iowa State University Press, Ames, Iowa, 1963.

13. Roberts, S. D., "The Development of a Self-Adaptive Forecasting Techniques," Unpublished MSIE Thesis, Purdue University, 1966.

14. Spencer, M. H., Clark, C. G., and Hoguet, P. W., *Business and Economic Forecasting*, Richard D. Irwin, Inc., Homewood, Illinois, 1961.

15. Spendley, W., Hext, G. R., and Himsworth, F. R., "Sequential Application of Simplex Designs in Optimization and Evolutionary Operation," *Technometrics*, Vol. 4, No. 4 (November, 1962).

16. Siegal, S., *Nonparametric Statistics*, McGraw-Hill, New York, 1956.

17. Winters, P. R., "Forecasting Sales by Exponentially Weighted Moving Averages," *Management Science* (April, 1960), pp. 324-342.